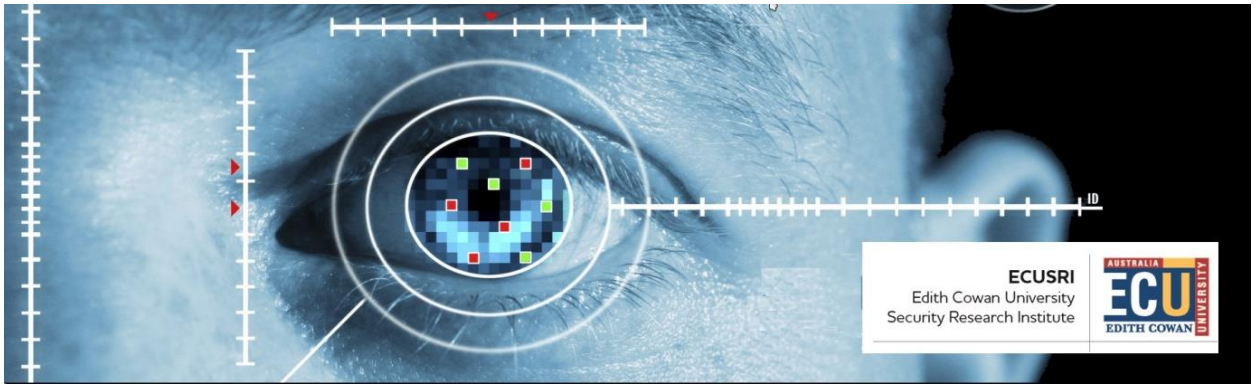
2016

The Proceedings of 14th Australian Digital Forensics Conference, 5-6 December 2016, Edith Cowan University, Perth, Australia

Craig Valli
Security Research Institute, Edith Cowan University, c.valli@ecu.edu.au



The Proceedings of

14th Australian Digital Forensics Conference

5-6 December 2016

Edith Cowan University, Perth, Australia

AUSTRALIA ECU EDITH COWAN UNIVERSITY

Proceedings of the
14th Australian Digital Forensics Conference

Published By
Security Research Institute
Edith Cowan University

Edited By
Professor Craig Valli
Security Research Institute
Edith Cowan University

ISBN 978-0-9945502-9-3

CRICOS Institution Provider Code 00279B

Sponsors

ECUSRI
Edith Cowan University
Security Research Institute

AUSTRALIA ECU EDITH COWAN UNIVERSITY

CISCO

Secure Systems

Supporters

AUSTRALIAN SECURITY MAGAZINE

acs AUSTRALIAN COMPUTER SOCIETY

Australian and New Zealand FORENSIC SCIENCE SOCIETY

Conference Foreword

This is the fifth year that the Australian Digital Forensics Conference has been held under the banner of the Security Research Institute, which is in part due to the success of the security conference program at ECU. As with previous years, the conference continues to see a quality papers with a number from local and international authors. 11 papers were submitted and following a double blind peer review process, 8 were accepted for final presentation and publication.

Conferences such as these are simply not possible without willing volunteers who follow through with the commitment they have initially made, and I would like to take this opportunity to thank the conference committee for their tireless efforts in this regard. These efforts have included but not been limited to the reviewing and editing of the conference papers, and helping with the planning, organisation and execution of the conference. Particular thanks go to those international reviewers who took the time to review papers for the conference, irrespective of the fact that they are unable to attend this year.

To our sponsors and supporters a vote of thanks for both the financial and moral support provided to the conference. Finally, to the student volunteers and staff of the ECU Security Research Institute, your efforts as always are appreciated and invaluable.

Yours sincerely,

Conference Chair
Professor Craig Valli
Director, Security Research Institute

Congress Organising Committee:
Congress Chair: Professor Craig Valli

Committee Members:	Dr Zubair Baig
	Mr David Cook
	Mr Michael Crowley
	Mr Peter Hannay
	Professor Bill Hutchinson
	Dr Ahmed Ibrahim
	Dr Mike Johnstone
	Mr Patryk Szewczyk
	Mr Krishnun Sansurooah
	Associate Professor Andrew Woodward
Congress Coordinator:	Ms Emma Burke

Table of Contents

DETECTING AND TRACING SLOW ATTACKS ON MOBILE PHONE USER SERVICE

Brian Cusack, Zhuang Tian
Digital Forensic Research Laboratories, AUT
brian.cusack@aut.ac.nz, zhuang_tian@hotmail.com

Abstract

The lower bandwidth of mobile devices has until recently filtered the range of attacks on the Internet. However, recent research shows that DOS and DDOS attacks, worms and viruses, and a whole range of social engineering attacks are impacting on broadband smartphone users. In our research we have developed a metric-based system to detect the traditional slow attacks that can be effective using limited resources, and then employed combinations of Internet trace back techniques to identify sources of attacks. Our research question asked: What defence mechanisms are effective? We critically evaluate the available literature to appraise the current state of the problem area and then propose an innovative solution for the detection and investigation of attacks.

Keywords: Slow Attacks, Detection, Trace back, Mobile, Communications

INTRODUCTION

There have been known security incidents of DDoS involving Mobile devices. For instance, September 2015, researchers from CloudFlare reported that a DDoS attack peaked at over 275,000 HTTP requests per second and resulted in 4.5 billion hits on the targeted website. This was blamed on a malicious advertising that compromised up to 650,000 Smartphones (Murdock, 2015, p.1). An update (2016) notes that these attacks spike during the weekend, they are very large in size, and that these attacks are no longer targeted only at high profile websites but also at mobile services. 3G technologies use IP technologies for control and transport; and, require cross network service collaborations, multi-vendor, and a multi-domain environment in order to gratify a wide variety of needs. This relationship requires Internet-based data and data from the cellular network in order to provide services to wireless users (Kotapati, et al., 2005, p.631). Bailey et al. (2009) reported that smart devices were responsible for generating 14 times more traffic than a non-smart device. As a result, the cellular networks have made tremendous improvements in order to meet the demands for increased bandwidth and communication requirements (Anstee et al., 2013). According to Farina et al. (2014), the 4G connections are responsible for generating six times more traffic than non 4G connections. However, globally, mobile data traffic reached 3.7 Exabyte per month in 2015; making mobile data traffic grow 4,000-fold over the past 10 years and almost 400-million-fold over the past 15 years. Smart devices represented 36 percent of mobile device connections globally in 2015. This accounted for 89 percent of mobile data traffic in which 55 percent was mobile video traffic. Consequently our paper acknowledges the trends but addresses the traditional slow attack that works with all mobile devices of any bandwidth (Farina et al., 2016).

The new venture between the two different technologies introduces new vulnerabilities and exposes the users on the cellular network to a range of additional risks across the new surface (Ricciato, et al., 2010, p.553).The introduction and growth of usages of technologies such as 4G/LTE and its high bandwidth has increased the pervasive nature of access points to the network. It is therefore evident mobile devices constitute not only a new target of an attack but also it has the capability to execute an attack (Farina, et al., 2016, p.269). Contact lists stored on mobile devices can be used to spread malware and infect other devices (Plohmann, et al., 2011, p.133). A DoS/DDoS attack has evolved from flooding strategies to low bandwidth tactics that employ slow techniques and can operate in all bandwidths. The purpose of this Slow DoS techniques is to lower the amount of bandwidth and resources that are required to execute an attack. The slow techniques have been adopted and used against devices such as mobile phones and game stations (Cambiaso, et al., 2012, p.195). While most of the packets sent to the target node in a flooding DoS attack may be useless but, in a low-rate attack, almost all of the packets play a role in the success of the attack. Therefore, the low-rate DoS will force the victim to process only the attack packets. There is not yet an effective tool to address an efficient detection method in relation to slow-rate DoS (p.197).

In this paper, we present an innovative technique to detect this kind of attack on mobile devices and also the use of multiple digital forensics methods to trace back the attack to its origin through the internet. The paper is designed to define slow attacks and to demonstrate our detection metric. We then present a flow diagram for investigation of slow attacks. The discussion on trace back reviews some of the previous and current literature in the field and draws the conclusion that two techniques working together are better than one on its own. We conclude by discussing

these claims and suggesting that slow attacks can be detected and managed to prevent loss of service to mobile phone users.

LOW RESOURCE ATTACKS

Low resource attacks rely on drip feeding malicious packets into a system. These techniques are often described as being slow because they are in contrast to flooding which rushes multiple packets through high bandwidth connections (Gilad and Herzberg, 2012). In contrast a low resource attack slowly drips malicious packets into a system such as a mobile phone where the victim processes every packet. Figure 1 shows the protocols which are open for exploitation by low resource slow attacks. The key element in each protocol is the ability for the attacker to slow down the attack packet by packet and to exploit the protocol mechanisms. Slow HTTP attacks, for example, rely on the fact that the HTTP protocol, by design, requires requests to be completely received by the server before they are processed. If an http request is not complete, or if the transfer rate is very low, the server keeps its resources busy waiting for the rest of the data. If the server keeps too many resources busy, this creates a denial of service.

Table 1. Protocols for slow attack.

Protocols	HTTP, HTTPS, ICMP, TCP, UDP, SYN, IRC
Attributes	Time interval

These types of attack are easy to execute because a single machine is able to establish thousands of connections to a server and generate thousands of unfinished HTTP requests in a very short period of time using minimal bandwidth. Due to implementation differences among various HTTP servers, two main attack vectors exist (Cambiaso, et al., 2012):

- **Slowloris**: Slowing down HTTP headers, making the server wait for the final CRLF, which indicates the end of the headers section; and,
- **Slow POST**: Slowing down the HTTP message body, making the server wait until all content arrives according to the Content-Length header; or until the final CRLF arrives, where if HTTP 1.1 is being used and no Content-Length was declared.

These attacks can just look like requests that are taking a long time, so it's hard to detect and prevent them by using traditional anti-DoS tools. In low resource conditions these attacks are effective because it does not require a large number of packets to create the effect. A defence against such an attack can be made by the following actions:

- Reject / drop connections with HTTP methods not supported by the URL.
- Limit the header and message body to a minimal reasonable length. Set tighter URL-specific limits as appropriate for every resource that accepts a message body.
- Set an absolute connection timeout nearing in mind that if the timeout is too short, you risk dropping legitimate slow connections; and if it's too long, you don't get any protection from attacks. A timeout value slightly greater than median lifetime of connections should satisfy most of the legitimate clients.
- The backlog of pending connections allows the server to hold connections it's not ready to accept, and this allows it to withstand a larger slow HTTP attack, as well as gives legitimate users a chance to be served under high load. However, a large backlog also prolongs the attack, since it backlogs all connection requests regardless of whether they're legitimate. If the server supports a backlog, make it reasonably large so your HTTP server can handle a small attack.
- Define the minimum incoming data rate, and drop connections that are slower than that rate. Care must be taken not to set the minimum too low, or you risk dropping legitimate connections.

However, these actions provide some protection but they do not signal a slow attack is taking place – which we address by innovation in the next section.

DETECTING ATTACKS

Distributed Denial of Service (DDoS) is simple but a very powerful technique of attack that disrupts service (Hadiks et al., 2014). The recent rapid proliferation and development of mobile technologies has also led to the exploitation for service disruption (Stafford and Urbaczewski, 2004, p.292). New techniques have also been developed to exploit the capacities and the characteristics of the service. This is where the DoS attack known as Low-rate DoS/DDoS attacks has evolved (Wang et al., 2007). Low-rate DDoS sends attack traffic periodically to the target device which makes it hard to detect amongst the normal traffic (Cambiaso et al., 2012). Various techniques for detection of the traditional flooding DDoS has been discussed in the literature. This section is designed to define and propose the

use of the distance based similarity metric to detect a Low-rate DDoS attack. The metric has been used in other contexts by Riccardio et al. (2010) who proposed a Similarity of Attack Intentions (SAI) to estimate the similarity of cybercrime intentions for network forensics. Another study found that using self-similarity algorithm to detect flooding DDoS attacks (Yu, 2014b). It has been used as the method for link predictions that compare one data set with another. For instance, *x* and *y* is assigned a score S_{xy} which can be defined as proximity or similarity between *x* and *y* (Yu, 2014a). The Similarity metric can be used in a more skilled approach such as using node attributes to define their similarity (Yu, 2014c). Similarity has been used also to evaluate distances between nodes. The shorter the path between nodes, the more similar they are (Snoeren et al., 2012). Distance based similarity metric is employed in this study to evaluate the similarity of the previous log file against the current log file in order to determine if a DDoS attack has occurred. Hadicks et al. (2014) argued that defining the problem will be the best way to fully understand the nature of the problem and what to match, i.e., what are the features to be used in matching; what are the constraints we have to consider; how to match, i.e., the matching process for achieving a consistent match; how to evaluate the match, i.e., define the similarity measure (p.3). The challenge in slow or low-rate DoS/DDoS attack is to map the proximity between attacking packets for identification, and then to initiate trace back methods based on the identified packets.

To evaluate the similarity between two different objects *x* and *y*, a distance metric known as Euclidean Distance (*EU*) is used. This metric can also be generalized into n-dimensions points, such that a={x_1, x_2, … x_n} and b={y_1, y_2, …, y_n}. In this case, n-dimensions *EU* metric is defined as:

$$EU\ (a,b) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n}(X_i - Y_i)^2}$$

To apply this metric to a Web server or to a mobile device then the log files have to be isolated for analysis. Let L_1 and L_2 be the existing log file and the current log file, respectively. Let x_i represent each protocol used in the existing log and y_i represent the protocol used in the current log, where $i=\{1, 2, ..., n\}$ and *n* is the total number of protocols where, L_1 ={x_1, x_2, …, x_n} and L_2 ={y_1, y_2, …, y_n}. For computational purposes the Euclidean distance can be normalized into a distance based similarity as follow: $S = \frac{1}{1+EU(L_1,L_2)}$

The normalized *EU* delivers a value in between 0 and 1, where a value of 1 means that the two objects are identical and a value other than 1 means that the two objects are not identical. Consequently the analysis focuses upon the discrimination between two known log files. The differential will indicate changes that can inform the alert of an attack. For the DDoS attack various protocols can be engaged in an attack (see figure 1). In order to detect an attack, the similarity between the existing and the current log files are ranked. In doing so, the Euclidean distance between L_1 and L_2 is calculated by using the first equation and then the similarity can be ranked based on the second equation. Table 1 provides a worked example of the detection metrics being applied to a sample set of mobile attack data that was downloaded from the Internet to illustrate the use of the detection system. It is a simple case of calculating the distance based similarity of various protocols that were used in the attack. The sample data was taken from a live attack and then processed. Once the attack has been detected, the protocol that was engaged in the

Table 2. A simple case of distance based similarity ranking.

Protocols	HTTP	HTTPS	ICMP	TCP	UDP	SYN	IRC
L_1	x_1=1000	x_2=800	x_3=600	x_4=2000	x_5=5000	x_6=6000	x_7=200
L_2	y_1=21000	y_2=1000	y_3=600	y_4=3000	y_5=7000	y_6=1000	y_7=500
$EU(L_1, L_2)$	***20000***	200	0	1000	2000	5000	300
S	***0.00005***	0.004	1	0.001	0.0005	0.0002	0.03

attack needs to be identified. This data in Table 2 shows in the S row that only one ICMP sample was the same. The variations in the other protocols indicates that an attack is occurring through them.

INVESTIGATING SLOW ATTACKS

Mobile forensics is defined as the science of recovering digital evidences from a mobile device under forensically sound conditions using accepted methods (Mumba and Venter, 2014, p.4). Mobile forensics investigation process consist of 15 phases that are divided into three main processes. The initialization process, the acquisition processes

and the investigative processes. (Omeleze and Venter, 2013, p.5). The investigative processes consists of six processes. The Potential digital evidence acquisition, digital evidence examination and analysis, digital evidence interpretation, reporting, presentation and investigation closure (Mumba and Venter, 2014, p.4). The processes employed in this study only concern the examination and analysis phase. The results will be used to determine a slow attack first and then initiate trace back the potential location of the attacker (Curran, et al., 2010; Omeleze et al., 2013). These processes were designed not only to eliminate the irrelevant data but assure the admissibility of the evidence in the court of law (Jansen, et al., 2007). The mobile forensics analysis process as illustrated in figure 1 starts with the data acquired from the victim's device. The data is used together with the reports from the similarity distance detection metric they can be calculated as a continuous live process or from previous log data.

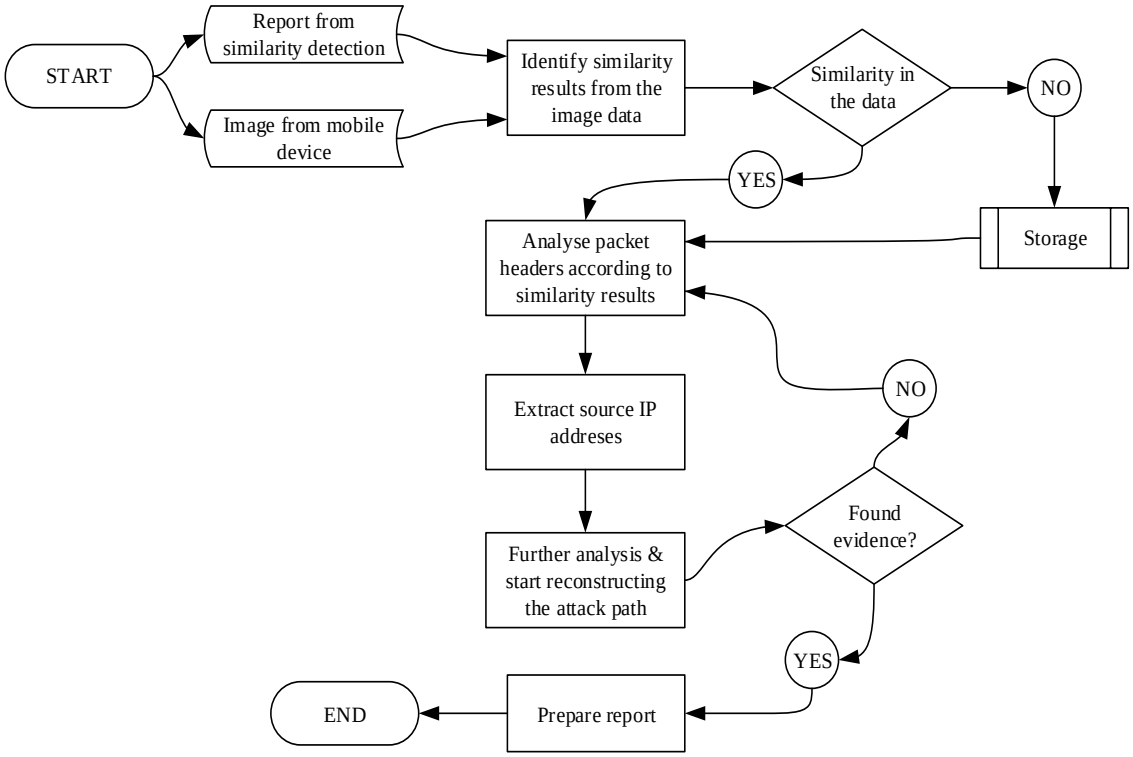


Figure 1. Mobile Low-rate DDoS forensics investigation process

Figure 1 assumes the common forensic soundness criteria are applied to correlate practice management by answering the questions: What meaning can be extracted from the evidence? What are the potential error factors? What are the training requirements for forensic practitioners? The model was developed to systematise processes while we investigated a number of slow attacks.

TRACEBACK

Traditional trace back methods are well-established for the Internet and broadband devices. The evidence presented in the introduction to this paper suggests that the majority of mobile phones are not broadband but the broadband phones produced most of the traffic. In addition all phones have some form of connectivity to the Internet. Therefore it is our argument that we can evaluate the traditional trace back methods and select the ones that are most appropriate for tracing back slow attacks. One reason that spoofing is often facilitated in these and other DoS or DDoS attacks is that it allows evasion of filters and quotas based on sender IP address, making tracing attackers harder (Devasundaram et al., 2006). Yu (2014b) reinforces that tracking back to the attack origin in DDoS attacks is a difficult and non-trivial problem due to the following reasons. Firstly, it is easy to forge or modify IP address (e.g. IP spoofing). Secondly, the stateless nature of IP routing, where routers normally know only the next hop for forwarding a packet instead of the entire end to end path taken by each packet, makes IP traceback even harder. Moreover, the Internet was originally designed for fast file sharing in a trusted environment and the network security was less important than communications, as it was a secondary consideration. Routers do not verify the source address of IP packets and the entire routing table is constructed on a trust basis. However, there are methodologies they can trace back to the last router from single packets. These methods can also be applied to trace back slow attacks, packet by packet (Goodrich, 2008).

A number of trace back methodologies can be rejected because they are impractical or too costly to implement. With slow attacks we are dealing with single packets or periodic clustered dispersions which are drip fed into a system to compromise devices. The metric can detect these time based malicious packets and once an alert sounded, trace back methods can be employed. A slow Denial of Service (DoS) attack can be characterized as an attack with the purpose of preventing legitimate users from using some specific network utilities such as a website, web service or computer system (Hadiks et al., 2014). It is also a coordinated attack on the availability of the service of a given target system or network. It can be launched indirectly through many compromised computing systems. The websites or other mobile devices used to launch the attack are often called the 'secondary victims' (Izaddoost et al., 2007). The use of secondary victims in a slow DDoS attack provides an attacker with the ability to launch a much larger and more disruptive attack than a slow DoS attack while remaining anonymous since the secondary victims actually complete the attack, and hence make it more difficult for the digital forensic investigator (DFI) to track down the original attacker. In general, there are two types of standard attacks (Leavitt, 2005): direct and reflector attacks. In a direct attack, an attacker sends attack packets directly towards the victims. Attack packets can be any of the protocols in figure 1. (Kumar et al., 2011). In each attack on a mobile device a variety of networks are being used. The first instance it may be a Wi-Fi connection, or a direct cellular signal, or any other wireless protocol. Although it may be theoretically possible to trace back in IP address and practice there are too many mediating factors, including spoofing, dynamic IP, and other obfuscations. However at the packet level the packets carry information regarding the pathway they have taken. For example the ICMP protocol can hold information regarding at least the last two or three services it transacted through routers. Hence, in particular for mobile devices connected to a wireless router trace back progress can be made.

Table 3. Comparison of Traceback Methods

Traceback Method	Hop Count Filtering	ICMP	Logging	Marking	Marking & Logging	TTL & Marking	FDDA
ISP Involvement	None	Low	Moderate	Low	None	None	None
No. of Attack Packets needed for traceback	1	Very Large	1	Very Large	1	Very Large	large
Processing Overhead	Very Low	Low	Low	Low	Very Low	Low	High
Storage	Very Low	Low	Low	High	High	High	High
Ease of Implementation	Yes	Yes	Yes	No	No	No	No
Scalability	Highest	High	Fair	High	High	Highest	Highest
Bandwidth Overhead	None	Low	None	None	None	High	High
No. of functions needed to implement	3	2	3	2	5	5	6
Ability to handle major DDOS attack	Yes	Yes	Yes	Poor	Yes	Yes	Yes
Classification	IDS Based	Proactive	IDS Based	Proactive	IDS Based	Proactive	IDS Based
OSI Model Layer and Protocols	IP, Network Layer	ICMP, Network Layer	IP, Network Layer	IP, Network Layer	IP, Network Layer	IP, Network Layer	IP, Network Layer

The intermediate routers (routers between the source and destination) will generate a special ICMP packet according to the probability of 1 out of 20,000 once it receives an IP packet. The ICMP packet will be sent to either the source or the destination host with equal probability. The router's path information is stored in the ICMP packet and is collected and analyzed at the destination host. With forward or back link information, two routers can be identified in the path; while with two links of information, three routers can be identified. Because only partial path information is contained in the ICMP packet, it will be extremely difficult to identify several attack paths under a flooding DDoS attack but for slow attacks the method is much more effective (Kumar et al., 2010). To consistently construct the full or partial path, the destination host has to match the original IP packet with its corresponding ICMP packet, and this can be difficult. However, if the ICMP method is used in conjunction with the hop count method then some of the obstacles to tracing back to the source of the malicious packets can be overcome. The basic idea of hop count filtering method is to identify spoofed IP packets by using the source OS address and the hop count value in the IP packet and the filter of the spoofed IP packet under DoS and DDoS attack. The rational is that most of spoofed IP packets do not carry hop count values that are consistent with the IP address being spoofed at victim's device. Hence, an IP-to-hop-count (IP2HC) mapping table is built by the use of our metric during operations to distinguish between malicious and normal traffic. The simulation results show that close to 90% of spoofed traffic was identified (Paxson, 2001; Plohmann et al., 2011). Once an accurate IP2HC mapping table is built, the inspection algorithm checks the source IP address and the final time-to-live (TTL) value in each packet. The hop count method is not precisely an IP traceback method, since it cannot accurately pin point the attacking origin. It can only give a list of possible routers associating with an attacking origin. If all of the routers on the list form a circle, then the victim is the center and the hop count distance is the radius. Coupled with the ICMP analysis these are the most effective ways to trace back a slow DoS/DDoS attack (Smoeren et al., 2004).

CONCLUSION

Detecting and tracing slow attacks on mobile phone user services is possible when combinations of methodologies are employed. We have demonstrated using dummy attack data from the web (Table 2) that our metric will detect and alert a slow denial of service attack from any of the protocols in Table 1. The review of trace back methodologies shows that many are not useful for slow attack but a combination of ICMP and the hop count methodologies can be effective by simply focusing on the packets. Disruption and other attacks will continue to grow on mobile devices of any bandwidth. Consequently, further research is required into detection, protection, and trace back methodologies in order to secure services.

REFERENCES

Anstee, D., Bowen, P., Chui, C. F., & Sockrider, G. (2016). Worldwide infrastructure security report. Special Repot: Arbor Networks, 11(1), 1-120.

Bailey, M., Cooke, E., Jahanian, F., Xu, Y., & Karir, M. (2009). A Survey of Botnet Technology and Defenses. Proceedings of the CATCH '09. Cybersecurity Applications & Technology Conference For Homeland Security (pp.299-304). Washington, DC: IEEE.

Cambiaso, E., Papaleo, G., & Aiello, M. (2012). Taxonomy of Slow DoS Attacks to Web Applications. In S. M. Thampi, A. Y. Zomaya, T. Strufe, J. M. Alcaraz Calero, & T. Thomas (Eds.). Proceedings of the International Conference, SNDS 2012 on Recent Trends in Computer Networks and Distributed Systems Security (pp. 195-204). Trivandrum: Springer.

Curran, K., Robinson, A., Peacocke, S., & Cassidy, S. (2010). Mobile Phone Forensic Analysis. International Journal of Digital Crime and Forensics, 2(2), 1-11.

Devasundaram, S. (2006).Performance evaluation of a TTL-based dynamic marking scheme in IP traceback. Akron, OH, USA: University of Akron.

Farina, P., Cambiaso, E., Papaleo, G., & Aiello, M. (2014). Mobile Botnets development: issues and solutions. International Journal of Future Computer and Communication, 3(6), 385-390.

Farina, P., Cambiaso, E., Papaleo, G., & Aiello, M. (2016). Are mobile botnets a possible threat? The case of SlowBot Net. Computers & Security, 58, 268-283.

Gilad, Y., Herzberg, A.: LOT: A defense against IP spoofing and flooding attacks. ACM Transactions on Information and System Security, 15 (2), 6 (2012)

Goodrich, M. T. (2008). Probabilistic Packet Marking for Large-Scale IP Traceback. IEEE/ACM Transactions on Networking, 16(1), 15-24.

Hadiks, A., Chen, Y., Li, F., & Liu, B. (2014). A study of stealthy denial-of-service attacks in Wi-Fi direct device-to-device networks. Proceedings of the 2014 IEEE 11th Conference on the Consumer Communications and Networking Conference (CCNC) (pp. 507-508). Las Vegas, NV: IEEE.

Izaddoost, A., Othman, M, Rasid, M.: Accurate ICMP traceback model under DoS/DDoS attack. ADCOM '07 Proceedings of the 15th International Conference on Advanced Computing and Communications (pp. 441-446). Washington, DC, USA: IEEE Computer Society (2007)

Jansen, W., & Ayers, R. (2007). Guidelines on Cell Phone Forensics: Recommendations of the National Institute of Standards and Technology. NIST: Special Publication 800-101, 1(1), 1-104.

Kumar, B., Kumar, P., Sukanesh,: Hop count based packet processing approach to counter DDoS attacks . International Conference on Recent Trends in Information, Telecommunication and Computing (ITC) (pp. 271-273). Kochi, Kerala, India: IEEE (2010)

Kumar, K., Sngal, A., Bhandari, A.: Traceback techniques against DDoS attacks: A comprehensive review. 2011 2nd International Conference on Computer and Communication Technology (ICCCT) (pp. 491-498). Allahabad, India: IEEE (2011)

Leavitt, N. (2005). Mobile phones: the next frontier for hackers? Computer, 38(4), 20-23.

Mirkovic, J., & Reiher, P. (2004). A taxonomy of DDoS attack and DDoS defense mechanisms. ACM SIGCOMM Computer Communication Review, 34(2), 39-53.

Mumba, E. R., & Venter, H. S. (2014). Mobile forensics using the harmonised digital forensic investigation process. Proceedings of the ISSA 2014 Conference on Information Security for South Africa (pp. 1-10). Johannesburg: IEEE.

Murdock, J. (2015). 650,000 Chinese smartphones used to launch ad network DDoS attack. Incisive Business Media, 1(1), 1-3.

Omeleze, S., & Venter, H. S. (2013). Testing the harmonised digital forensic investigation process model-using an Android mobile phone. Proceedings of the ISSA Conference on Information Security for South Africa, (pp.1-8). Johannesburg: IEEE.

Paxson, V.: An analysis of using reflectors for distributed denial-of-service attacks. ACM SIGCOMM Computer Communication Review, 31 (3), 38-47 (2001)

Plohmann, D., Gerhards-Padilla, E., & Leder, F. (2011). Botnets: Detection, Measurement, Disinfection & Defence. European Network and Information Security Agency (ENISA), 1(1), 1-153.

Ricciato, F., Coluccia, A., & D'Alconzo, A. (2010). A review of DoS attack models for 3G cellular networks from a system-design perspective. Computer Communications, 33(5), 551-558.

Snoeren, A., Partridge, C., Sanchez, L., Jones, S., Tchakountio, F., Schwartz, B., Kent, S., Strayer, W. (2012). Single-packet IP traceback. IEEE/ACM Transactions on Networking, 10 (6), 721-734 .

Stafford, T. F., & Urbaczewski, A. (2004). Spyware: The ghost in the machine. The Communications of the Association for Information Systems, 14(1), 291-306.

Wang, H., Jin, C., Shin, K.: Defense against spoofed IP traffic using hop-count filtering. IEEE/ACM Transactions on Networking, 15 (1), 40-53 (2007)

Yu, S. (2014a). Attack Source Traceback. In Distributed Denial of Service Attack and Defense (pp. 55-75). NY: Springer.

Yu, S. (2014b). An Overview of DDoS Attacks. In Distributed Denial of Service Attack and Defense (pp. 1-14). NY: Springer

Yu, S. (2014c). DDoS Attack Detection. In Distributed Denial of Service Attack and Defense (pp. 31-53). New York, NY: Springer

MEMORY FORENSIC DATA RECOVERY UTILISING RAM COOLING METHODS

Kedar Gupta, Alastair Nisbet
Security & Forensic Research Group, Auckland University of Technology
Auckland, New Zealand
kedargupta@msn.com, alastair.nisbet@aut.ac.nz

Abstract

Forensic investigations of digital devices is generally conducted on a seized device in a secure environment. This usually necessitates powering down the device and taking an image of the hard drive or semi-permanent storage in the case of solid state technology. Guidelines for forensic investigations of computers advise that the computer should be shut down by removing the power supply and thereby maintaining the hard disk in the state it was in whilst running. However, valuable forensic evidence often exists in the volatile memory which is lost when this process is followed. The issues of locked accounts on running computers and encrypted files present particular difficulties for forensic investigators who wish to capture a forensic image of the RAM. This research involves freezing RAM removed from a running computer so that it can later be reinserted into an unlocked computer allowing for a forensic image of the RAM to be captured. Three different methods of cooling the RAM are compared, along with varying delays in RAM reinsertion. The results provide a guideline for forensic investigators on how the issues with locked accounts and encryption may be overcome to record this valuable evidence that is otherwise lost.

Keywords
information security, RAM, forensics

INTRODUCTION

The forensic investigation of digital devices usually involves an investigator examining a cell phone, tablet of computer hard disk in a sterile environment. Guidelines for forensic investigators often advise that if a computer is running, it should be switched off at the power source to preserve the hard drive in its current state (National Institute of Justice 2008). One authoritative guide for forensic investigations suggested the need to preserve all evidence yet then suggested that power should be removed immediately from a running computer. This was corrected in a later version where a greater recognition of the value of evidence in the memory was recognised. (Association of Chief Police Officers (ACPO) 2012). Shutting down the computer using the normal method will inevitably cause data to be written to the hard drive possibly overwriting valuable forensic evidence from deleted files (Vidas 2007). Whilst this method focusses on the potential forensic evidence on a hard drive, the Random Access Memory (RAM) is often overlooked as a source of forensic evidence. RAM may contain highly valuable forensic data such as web browsing history, programs that have been recently run or werecurrently running and usernames and passwords in unencrypted form (Halderman, Schoen et al. 2009). This highly valuable and useful information will be lost the instant power is switched off to the computer (Simon and Slay 2009).

It is fairly trivial to insert a USB drive with the required software to download the contents of RAM on to an external drive for later analysis. Whilst often this is not done, the fact that RAM may contain several gigabytes of valuable forensic evidence means that evidence is lost because the forensic investigator has not considered this area to be valuable to an investigation. One problem that may arise for an investigator is that of a computer that has been locked by the user and requires a password to unlock the user account. Whilst this may not present a problem in recovering hard drive evidence, the shutting down of the computer will permanently delete the volatile evidence in memory. Another significant issue that has surfaced in recent years for forensic investigators is the increasing use of cryptographic tools to encrypt files or entire hard drives. The cryptography is generally robust enough to withstand all attempts at finding the encryption key required to unlock this evidence. A securely encrypted drive means that evidence can be seized but that it simply cannot be read rendering it useless.

This research examines the area of RAM forensics as an important source of digital forensic evidence for investigators. Methods of freezing RAM are utilised to compare the retention of data within the RAM utilising three different methods. Results are then compared so that guidelines for forensic investigators are presented allowing recovery of RAM contents using these methods.

STATE OF THE ART

The physical address space consists of all valid physical memory addresses that are readable by the CPU (Intel Corporation 2013). The Memory Management Unit (MMU) acts as an intermediary between the operating system and the memory so that any attempt at acquiring the contents of RAM will require the utilised software to overcome restrictions created by the MMU. Whilst the utilisation of address spaces may vary across architectures and operating systems, the address spaces in RAM are all potentially accessible through specialised software.

The layout of the physical address space in RAM is finalised upon completion of the copying of firmware code from ROM to RAM (Salihun 2012). Once this is complete, the RAM is ready for use by the operating system. Some forensic guidelines suggest photographing a computer screen and then switching the computer off at the power supply but forensic investigators are often aware of the valuable evidence contained in RAM and suggest performing a memory dump prior to powering off the system (Ligh, Case et al. 2014). Whilst a memory dump of RAM should provide all contents of RAM, it is possible that memory smear may occur because of the time taken to acquire the RAM contents (Vömel and Freiling 2012). Memory smear occurs when memory is copied from RAM to a physical media and during this process some errors occur because of the time taken for the acquisition. This results in minor differences between the RAM contents and the RAM image so that whilst a 100% match of the contents is hoped for, often the image falls slightly below a 100% match.

To overcome this problem, a hardware-based approach using a Firewire connection can be utilised which provides more direct access to RAM, decreasing the error rate to the point that often a 100% match can be achieved (Huebner, Bem et al. 2007). If the Firewire approach is not feasible and the computer is running a Windows operating system, then 2 other methods may provide greater reliability than a memory dump (Ruff 2008). Firstly, intentionally crashing the operating system will cause the RAM contents to be copied to the hard disk meaning a forensic acquisition of the hard disk should include most, if not all of the contents of RAM. The drawback to this method on a Windows Operating System is that the same computer must be rebooted to allow memory to be rewritten to the RAM (Russinovich, Solomon et al. 2012). The second method is to utilise the hibernation feature which will copy the contents of RAM to the hard disk and then power down the computer.

Whilst Firewire may provide greater reliability in copying memory to external storage, and the other methods suggested may provide sufficient integrity of the data for Windows Operating Systems, the forensic investigator may be in a position where these options are not available. Additionally, these options won't work if the computer has been locked by the user so that access to the RAM is not available. This problem is common and whilst asking for the password may be an option, a legally savvy suspect will likely refuse to give the password. Furthermore, full disk encryption is becoming more commonplace with freely available software such as TrueCrypt that will encrypt all files on a disk rendering them unreadable unless the password is disclosed that will decrypt the files. One issue that may also present to investigators during the memory acquisition phase is a deliberate attempt to corrupt the contents of memory if a memory dump is attempted. Specialised software designed to corrupt the memory can be created and will run when the memory is dumped from the operating system (Milkovic 2012). This can be overcome if the anti-forensic can be disabled or the RAM can be removed and inserted into another, similar but clean computer.

In 2016, an experiment was run in freezing RAM using freezing spray (Bauer, Gruhn & Freiling, 2016). In this experiment the RAM was cooled insitu with the computer running. The focus on the experiments was to demonstrate the recovery of the data and how it could be unscrambled to produce an original image or file. Unlike the current experiments, it did not utilise cooling with liquid nitrogen or ice and so has a different focus than this research.

The following section describes the experiments in cooling RAM so that the forensic investigator can remove the ram yet maintain the integrity of the data for a sufficient period of time to permit forensic analysis of the contents.

EXPERIMENTAL DESIGN

The problem for forensic investigators of locked user accounts means that RAM cannot be examined because an unlocked computer is required to examine and download RAM. The computer is locked by the user selecting 'ctrl alt delete' which then requires that same user's password to be entered to unlock the computer. What is

therefore required from a forensic investigator is the ability to remove RAM from a locked computer without losing the integrity of the digital contents and to either place the RAM into an identical computer or to reboot the suspect computer into a state where it is not locked. The RAM can then be inserted into the unlocked computer and a memory dump performed. The first stage in the experiments was to boot a laptop computer to the Windows 7 operating system. This system was chosen as it remains one of the most common operating systems in organisations and for home users. Six actions were then performed to provide data in the 1 GB RAM that would be useful in an investigation.

Those actions were:

1. Log on to Windows 7 with the administrator username and password.
2. Open a text file and copied text to the clipboard.
3. Started an Apache HTTP web server and MySQL database.
4. Unlocked a TrueCrypt file container and opened several files from the container.
5. Executed several commands using a DOS prompt.
6. Launched a web browser and logged into a Facebook account.

The process for calculating the percentage of data recovered was to take an image of the RAM and then to lock the account on the computer. The power button was then held down to halt power to the motherboard and the RAM removed at the same time where it was immediately put through the cooling process. Once the full experiment was completed and a RAM image obtained, this image was then compared to the prior image and the percentage of data recovered calculated.

Whilst this gives an accurate percentage of the data recovered, it does so at a macro level so that if parts of files are recovered or even parts of passwords, this is not obvious from the percentages that are recovered. To ascertain at a micro level, a more manual examination is required and in the case of recovered passwords this was done to ensure that a full password or encryption key has been recovered. The parameters for the test bed are shown in table 1:

Table 1: Testbed Setup

Item	**Description**
RAM	1GB DDR3
Error Correction?	Non Error Correction RAM
CPU	Samsung 2.4GHz

The experiments with freezing RAM to recover its contents began with a baseline to compare the results against. The baseline simply involved removing RAM from a computer and inserting the RAM into the now unlocked and rebooted computer. The minimum time between removal of the RAM and insertion into the computer was 10 minutes, allowing ample time for the investigator to reboot the machine to a usable state.

The surface temperature of the RAM was measured at 35 degrees centigrade. As expected, the contents of RAM were deleted upon removal so that no usable forensic evidence could be recovered using this method.

Table 2 shows the experimental setup for the RAM utilising three different cooling methods. The ambient temperature of the RAM before cooling was 35° degrees centigrade.

Table 2: RAM Freezing Experimental Steps

Method	Time Delay	Temperature Cooled degrees C
Liquid Nitrogen	10mins, 1 hr, 2hrs	-196
Freezing Spray	10 mins	-40
ICE	10 mins	10

The next step was to utilise liquid nitrogen to freeze the RAM. This has been attempted by previous researchers with good results. The RAM was removed from the running but locked computer and immediately submerged in liquid nitrogen, almost instantly freezing the RAM and it was hoped that the contents would therefore remain. The surface temperature of the RAM was measured at -196 degrees centigrade. The experiment was repeated a number of times with varying longer delays in removing the RAM from the liquid nitrogen to see of entropy of the contents occurred. The experiment was then repeated but with a spray of freezing compressed air rather than liquid nitrogen. The freezing by compressed air was a simple process and could easily be performed by a forensic investigator in the field as opposed to the practicality of carrying liquid nitrogen to an investigation scene. The surface temperature of the RAM was measured at -40 degrees centigrade. Finally, the third stage involved removing the RAM, placing it in an anti-static bag and burying it in ice cubes for 10 minutes. The surface temperature was measured at 10 degrees centigrade just immediately after removing it from the ice. Once cooled, the RAM was then reinserted into the same computer and the computer booted from a USB stick with DumpIT, a Linux memory dumping software programme installed on the stick which creates an image of the RAM. In the initial boot stages, the software immediately dumped the contents of the RAM onto a storage area on the USB stick where it remained for later analysis. The results of the RAM cooling are described below.

RESULTS

The purpose of the experiments was to ascertain whether freezing RAM with compressed air or ice would provide forensic evidence from the RAM that could be utilised in an investigation. It was ascertained that removing RAM and reinserting it 10 minutes later led to almost all contents being deleted. The liquid nitrogen provided results that led to almost a complete recovery of the RAM but was not practical for a forensic investigation in the field. Therefore, the freezing spray and submergence in ice were of particular interest as these methods are both feasible for an investigator. A can of freezing spray is easily obtainable for negligible cost and ice may be available if a forensic investigator has no freezing spray readably available. The following table shows the results in percentage of data recovered from RAM after 10 minutes.

Table 3: Percentage of forensic data recovered

Method	Temperature degrees Celsius	Percentage Recovered
No cooling	35	0.2%
Liquid Nitrogen	-196	99.81%
Freezing Spray	-40	96.45%
Ice	10	99.71%

The results indicated that whilst liquid nitrogen performed the best, submerging the RAM in ice was a very close second best with only 0.1% less data recovered. Freezing spray, whilst the most practical for a forensic investigator to carry in their toolkit, results in 96.45% data recovered. Whilst this result is significant, the 3.3% data not recovered compared to ice may be vital evidence that would be lost when choosing this method. Next, the results for submerging the RAM in liquid nitrogen were extended from previous experiments by allowing

for a much greater delay in submerging the RAM and removing it from the liquid nitrogen. The purpose of these further experiments were to simulate a forensic investigator who is equipped with a container of liquid nitrogen who may be able to simple drop the RAM into the container but may require a sterile and safe environment equipped with protective clothing to remove the RAM from the nitrogen. The time delay was therefore extended from 10 minutes to 1 hour and then repeated for a delay of 2.5 hours. The results are shown in table 4.

Table 4: Time delay with liquid nitrogen

Time delay	Percentage Recovered	Difference
10 minutes	99.81%	
1 hour	98.84%	-0.97%
2.5 hours	93.95%	-4.89%

It is clear from table 4 that even at -196% Celsius data is lost from RAM if there is a delay in reinserting the RAM into an unlocked computer. A delay of 1 hour results in a loss of over 1% more data than by utilising a freezing spray. If a delay of 2.5 hours occurs, then the loss exceeds the freezing spray by 2.5%. A delay of 1 hour may be unrealistic in many scenarios where the submergence of RAM would be followed by seizing of the computer and related activities, meaning that returning to a safely equipped laboratory may well take longer than 60 minutes. The delay of 2.5 hours is more realistic and shows a significant loss of valuable data. The rather low loss of 0.29% of data with ice indicates that this is the preferred method. It shows that loss of data is minimal and it is a very safe method, although somewhat cumbersome for the forensic investigator in the field. The preferred method may be to carry a can of compressed air as this is simple and fairly safe, but the forensic investigator must bear in mind that approximately 3.55% of data will be lost with a delay of only 10 minutes.

Finally, the data that was recovered from the RAM was identified. A USB drive with Linux software suitable for a RAM dump was used to store the RAM image. Among the data recovered was the expected information regarding user activities on the computer including web pages visited and processes that were started and stopped. Of particular interest was whether the TrueCrypt password could be recovered. It was found in all experiments that the 512 bit AES encryption key utilised by TrueCrypt was recovered allowing the encrypted files to be decrypted and read. This is particularly important for the forensic investigator because it is often the encrypted files that contain the most vital information for the forensic investigator. The recovery of the encryption key was possible in all experiments involving cooling of the RAM but with less than 100% of data recovered there was some luck involved. Any data that is lost may by bad luck involve the encryption key or other vital evidence and so the forensic investigator should bear this in mind when selecting which method to utilise. The greater the percentage of data recovered, the greater the likelihood that the most vital data will remain in the RAM image.

CONCLUSION

The purpose of a forensic investigation of a computer is to acquire, analyse and potentially use information in a criminal, civil or employee investigation. The more information that can be gleaned from the device, the higher the likelihood of a successful investigation. Traditional methods of preserving information from a running computer usually involve looking at the screen to ascertain running programs and then removing the power to the device to preserve the information on the hard disk. Often, the valuable forensic information in RAM is overlooked, or thought to be unobtainable because of a locked user account or encrypted files that prevent the investigator from reading the files unless the encryption key can be obtained. This is unlikely if the suspect's computer contains malicious or nefarious evidence of wrongdoing. The ability to secure the RAM image is therefore highly valuable to the forensic investigator.

This research provides evidence that RAM images can be obtained in these circumstances, and in the case of TrueCrypt encryption, the encryption key may be able to be recovered from the image of memory. This research provides a guideline for forensic investigators as to the best methods for a successful data recovery of RAM and provides evidence of the importance of time delays in the entropy of information in the case of liquid nitrogen freezing. The importance of memory forensics should not be overlooked by the forensic investigator and the

implementation of account locking or encryption of files or entire volumes does not present an insurmountable challenge to successfully acquiring a RAM image.

REFERENCES

Association of Chief Police Officers (ACPO) (2012). Good Practice Guide for Computer-Based Electronic Evidence. E-Crime Working Group.

Bauer, J. Gruhn, M. Freiling, F. (2016). "Lest we forget: Cold-boot attacks on scrambled DDR3 memory". Digital Investigation: 16, 65-74

Halderman, J. A., S. D. Schoen, et al. (2009). "Lest we remember: cold boot attacks on encryption keys." Communications of the ACM 52(5): 91-98.

Huebner, E., D. Bem, et al. (2007). "Persistent systems techniques in forensic acquisition of memory." Digital Investigation 4(3-4): 129-137.

Intel Corporation (2013). Intel 64 and IA-32 Architectures Software Developers' Manual. 3A: System Programming Guide, Part 1.

Ligh, M. H., A. Case, et al. (2014). The Art of Memory Forensics: Detecting Malware and Threats in Windows, Linux and Mac Memory, Wiley.

Milkovic, L. (2012). "Defeating Windows Memory Forensics." Retrieved June 1st 2016, from https://events.cc.de/congress/2013/Fahrplan/events/5301.en.html.

National Institute of Justice (2008). Crime Scene Investigation: A Guide for First Responders. U.S. Department of Justice. Washington, DC.

Ruff, N. (2008). "Windows Memory Forensics." Journal in Computer Virology 4(2): 83-100.

Russinovich, M. E., D. A. Solomon, et al. (2012). Windows Internals, Pearson Education.

Salihun, D. (2012). Disassembly Ninjutsu Uncovered (Uncovered series), A-List Publishing.

Simon, M. and J. Slay (2009). Enhancement of Forensic Computing Investigations through Memory Forensic Techniques. International Conference on Reliability and Security, ARES Fukuoka, Japan.

Vidas, T. (2007). "The Acquisition and Analysis of Random Access Memory." Journal of Digital Forensic Practice 1(4): 315-323.

Vömel, S. and F. C. Freiling (2012). "Correctness, Atomicity and Integrety: defining criteria for forensically-sound memory acquisition." Digital Investigation 9(2): 125-137.

IMPROVING FORENSIC SOFTWARE TOOL PERFORMANCE IN DETECTING FRAUD FOR FINANCIAL STATEMENTS

Brian Cusack, Tau'aho 'Ahokovi
AUT University, Christ's University in Pacific
brian.cusack@aut.ac.nz, tahokovi@bigpond.com

Abstract

The use of computer forensics is important for forensic accounting practice because most accounting information is in digital forms today. The access to evidence is increasingly more complex and in far greater volumes than in previous decades. The effective and efficient means of detecting fraud are required for the public to maintain their confidence in the reliability of accounting audit and the reputation of accounting firms. The software tools used by forensic accounting can be called into question. Many appear inadequate when faced with the complexity of fraud and there needs to be the development of automated and specialist problem-solving forensic software. In this paper we review the context of forensic accounting and the potential to develop improved support tools. The recommendation is for adopting financial ratio analysis as the basis for an improved fraud detection software.

Keywords
Financial Ratio Analysis, Forensic accounting, Forensic Software Tools, Fraud

INTRODUCTION

Before a potential fraud can be investigated, it must be detected. The process of fraud detection involves searching for symptoms that may indicate that fraud exists (Kovalerchuk et al., 2007, p.2). One of the most popular mathematical techniques which are currently used in forensic accounting, is that Benford law. The main issue with this technique is that it takes a relatively broad, rather than narrow approach to detecting fraud. As a result a lot of false signals can occur with the impact of time consumption and increase costs in investigations. Another technique known as relative size factor (RSF), is designed to detect any outliers or unusual data. However, any outlier can be just errors in data entry but not financial fraud. Panighahi (2006, p.3) stated that "RSF is simple to calculate but not an effective and efficient tool". An abundance of data creates both challenges and opportunities for the forensic accountant. A digital forensics software tool is employed to help deal with the big data problem and to speed up both the accuracy and the completion of investigations. The problem with computer based fraud detection in the field of forensic accounting is that there are significant differences in task performance and knowledge requirements for the completion of investigation. For example, computer forensics requires knowledge of computing systems, log files, graphics and other formats, and many other non-accounting knowledges. Similarly, data theft prevention and investigation requires database knowledge, computer security knowledge, encryption and computer systems (Albright, 2008, p.5).

In the field of forensic accounting, there are categories of investigation such as:

- *Data mining for fraud* Techniques and methodologies for discovering fraud in corporate databases;
- *Financial statement fraud:* Ratio analysis and other methods of finding financial statement manipulation
- *External information sources:* Information about perpetrator finances and other data, usually found in websites; and,
- *Computer forensics:* Investigating by sifting through computer hard drives and other information devices.

Each context has its own challenges and ways of investigation. As a result, digital forensic software tools have to declare capability before use. Otherwise a null or an erroneous result may be obtained from the tool as it is not capable in a particular area, and yet fraud exists within the dataset. Therefore, till testing is a critical issue that impacts potential evidential outcomes. Verification is required from approved by independent bodies and not just proprietary vendor's (Al Mutawa, et al., 2012, p.26). Researchers in the field suggested that it is vital for investigators to compare the suitability of forensic tools in relation to various application environments (Guo & Slay, 2010, p.297). Some have common views in analysis of large data (e.g. several terabytes) where tools should be able to efficiently and effectively handle the volumes (Yannikos et al., 2011, p.200). Kimmel et al., (2012) recommended that an in – depth assessment of the tool based on requirements and an evaluation of different packages and their functions within all available types of fraud pattern. What is missing from this

advice are statements about how software forensic tools will stay up-to-date and adapt themselves to new patterns of fraudulent activity (p.765).
In this paper we briefly review the problem area of relevancy for current digital forensic and accounting forensic software tools. We then propose a focus for the development of accounting forensic tools that takes ratio analysis and pattern recognition to be the foundational building block for the analysis. We then propose a fraud detection model that is based on the similarity metric for determining the similarity of patterns and the proximity of data to events. We conclude by recommending care in the use of current forensics software tools and emphasise the necessity of tools being able to not only cope with big data but also complexity.

THE PROBLEM AREA

Forensic accounting is the application of accounting knowledge and investigative skills to identify and document potential matters with legal implication for fraud and other financial crime (Houck et al., 2006, p.68). Forensic accounting is a developing area of specialisation in the field of accounting. Its main concerned with the detection and prevention of financial frauds and other forms of economic crime (Dhar and Sarkar, 2010, p.94). In the accounting field, there are a number of ongoing activities that, collectively, allow business owners or managers to access the information when they need it in order to make well informed decisions. In terms of business operation, this includes basic transactions such as purchases and sales, and marketing and strategic planning as well as summative information. Accounting refers to financial record keeping and data reporting that businesses engage in to meet legal requirements and keep the organisation's stakeholders informed of the organization's financial position at any point of the year. Accounting is defined as a systematic process of identifying, recording, measuring, classifying, verifying, summarizing, interpreting and communicating financial information. It reveals profit or loss for a given period, and the value and nature of a firm's assets, liabilities and owners' equity (Business Dictionary, 2016, p.1). However, the financial statement will only be as good as the journal entries (Haber, 2004, p.7). Missing or fraudulent journal entries will produce financial statement that is fraudulent (Basuhail, 2010, p.97).

Financial statements are prepared to present fair information about the financial position, operating performance of the organisation or the business. The international standards on auditing 240 stated that, fraud and error must be considered when auditing financial statements (Smith et al., 2008, p.18). Furthermore, the auditor must perform procedures to assist the entity in the detection of fraud (Council, 2013, p.13). As a result of the Enron and WorldCom failures (Reinstein and McMillan, 2004, p.956), the accounting arena had undergone fundamental changes to redress the shortcomings found in previous audit requirements. Therefore, a new market with a new class of accountants known as forensic accountants has emerged. The white collar crimes is in focus and occupational fraud. The Association of Certified Fraud Examiners (ACFE) estimates that occupational fraud losses cost organizations $994 billion annually (Davis et al., 2010, p.5; Crumbley et al., 2005, p.400). Forensic accounting and fraud examination are different but related. Forensic accounting work is done by accountants in anticipation of litigation and can include fraud, valuation, bankruptcy and a host of other professional services. Fraud examinations can be conducted by either accountants or non-accountants and refer only to anti-fraud matters (Wells, 2003, p.76).

DIGITAL FORENSIC SOFTWARE TOOLS

The amount of data stored in accounting files requires digital management. This means that relevant and appropriate software must be available to process the data and to extract the relevant information. Forensic investigation and the analysis of accounting data for fraudulent patterns has become more and more complex (Albano, et al., 2011, p.381). Digital forensics integrates the fields of computer science and law to investigate crime (Dezfouli, et al., 2012, p.186). For any digital evidence to be used in court, investigators must follow a proper set of procedures when collecting and analysing data from computer systems (Jansen & Ayers, 2007, p.6). Hence the interface between the investigator and the data is mediated by software but due to the differences in terms of the technologies, investigators will have to engage different methods and tools depending on the category of information involved (Albano, et al., 2011, p.381). Therefore, prior to acquiring data from a comprised device, extra cautions must be taken, standard procedures and base practises must be followed carefully. This is to avoid altering data in the process because digital data can be easily corrupted (Jansen & Ayers, 2007, p.45). The first challenge for forensic software tools is the bridging of the structures, protocols, and designs in which the data resides. The second challenge is then to be able to identify accurately patterns of conformance, compliance, and aberrations that may occur within the data (Ayers, 2007, p.1). Each tool has a different core set of features that are designed to deliver specific outcomes (NIST, 2013, p.4). Reliable digital forensic techniques are therefore important for prevention, detection, and investigation of electronic crime (Nissan, 2012, p.843). For example, in table 1 and analysis has been completed of different accounting forensic

and digital forensic tools that may be employed in an investigation. The analysis shows that each tool has a different capability and that some do overlap. However, an investigator has to be fully aware of these limitations before a tool is selected or used for forensic accounting purposes (Mohtasebi & Dehghantanha, 2013, 353). For most the integrity of the evidences and its admissibility in the court of law has to be preserved. As a result, it is critical for an investigator to know the reliability and accuracy of the tool (Kubi et al., 2011, p.2) and to select one's that are to be effective and efficient in the particular situation.

Table 1. Experiment results (Grispos, et al., 2011, p.30).

Item	Type	Logical Acquisition	Manual Examination	Physical Analyzer	Scalpel (configured)	Foremost (default)	Foremost (configured)	Simple File Carver	Phone Image Carver	WinHex (modified image)
1	docx	N	P	F	N	P	N	N	D	F
2	docx	N	P	F	N	P	N	N	D	F
3	rft	N	P	F	N	N	N	N	N	F
4	txt	N	P	F	N	N	N	N	N	F
5	xslx	N	P	F	N	P	N	N	N	F
6	pptx	N	P	F	N	P	N	N	N	F
7	pdf	N	P	F	F	F	P	D	P	F
8	pdf	N	P	F	D	D	D	D	D	F
9	jpg	F	P	F	D	F	D	D	D	F
10	jpg	F	P	F	D	F	D	D	D	F
11	jpg	F	P	F	D	F	D	D	N	F
12	jpg	F	P	F	D	F	D	D	N	F
13	jpg	F	P	F	D	F	D	D	D	F
14	mp3	F	P	F	P	N	P	N	D	F
15	wav	F	P	F	P	P	P	P	P	F
16	avi	N	P	F	D	D	D	N	D	F
17	wmv	N	P	F	P	P	P	P	P	F
18	mp4	F	P	F	D	N	D	N	N	F
19–23	Appointments	N	P	N	N	N	N	N	N	N
24–28	Contacts	F	P	N	N	N	N	N	N	N
29–30	Email Sent	N	P	P	F	N	F	F	N	N
31–32	Email Received	N	P	P	F	N	F	F	N	N
33–35	SMS Sent	N	P	P	N	N	N	N	N	N
36–38	SMS Received	N	P	F	N	N	N	N	N	N
39–43	Visited (IE)	N	P	P	F	N	F	F	F	P
44–50	Visited (Opera)	N	P	P	N	N	N	N	N	P
51	Favorite Websites	N	P	P	F	N	F	F	F	P
52–54	Call From	F	P	N	N	N	N	N	N	N
55–56	Call To	F	P	N	N	N	N	N	N	N
57–68	Deleted Files	N	N	D	N	N	N	N	N	D
69–70	Deleted Appointments	N	N	N	N	N	N	N	N	N
71–72	Deleted Contacts	N	N	N	N	N	N	N	N	N
73–74	Deleted Emails	N	N	N	N	N	N	N	N	N
75–77	Deleted SMS	N	N	N	N	N	N	N	N	N
78–79	Deleted Visited	N	N	N	N	N	N	N	N	N
80–82	Deleted Call Logs	N	N	N	N	N	N	N	N	N
Full		18	0	21	11	6	10	10	6	18
Partial		0	56	20	3	6	4	3	3	13
Detected		0	0	12	8	2	8	6	8	12
Not applicable		64	26	29	60	68	60	63	65	39

Keys that the authors used in Table 3.2 are F = Full, P = Partial, D = Detected and N = Not.

In spite of everything, all logical data can be acquired and analysed yet, tool developers seem to over claim their tool's support while the tool can only obtain some of the requirements. As a result, forensic tools should be evaluated based on their abilities and not the costs. For example, a tool may have better support for a particular brand of operating system or device (Morrissey, 2010, p.130). Many of the current software tools for forensic investigation may not be used in financial related investigation beyond the extraction of data because they do not discriminate sufficiently at the higher complexity levels. A forensic accountant requires detection of fraud in financial statement reports which is more than just extracting the report data. As a result, the abstraction of models is required to detect fraud in a financial report.

ACCOUNTING RATIOS

The initial design of the proposed model was developed based on the learning from the analysis of the literature, formulated to fill the gap, and avoid repetition. The reading and analysis gain sufficient abstraction that the modelling system would rest above the data level. The result is a tentative detection model for financial fraud that requires testing and validation. The initial design of the financial fraud detection model comprises of five financial ratios - *Return on Assets (ROA), Accounts Receivable (A/R) to sales ratio, Current Ratio, Total Asset*

Turnover and Inventory Turnover. The implication is that for a financial report - which brings together many segregated areas – financial ratios from each of the implicated areas are required comparison against the industry standard or a benchmark from the trend analysis ratios (Albano, et al., 2011). Models are an abstraction of a process to examine potential evidence, irrespective of the originality of the evidence (Peisert, et al., 2008, p.116). Forensic experts also believe that forensic investigation process models generalise an informal procedure to deliver a framework. That framework provides a detailed understanding of what each process is to do and not do. Jankun-Kelly, et al. (2007) explained that the model and framework provides an effective means to acquire information within the process. These processes are used to capture relevant aspects of the investigation (p.357). A breakdown of each ratio for building into the tool is as follows:

Return on Assets (ROA), Accounts Receivable (A/R) to sales ratio, Current Ratio, Total Asset Turnover and Inventory Turnover: If an organisation or business is subject to an audit, the auditors will review its accounts receivable in some detail. Accounts receivable is frequently the largest asset that a company has, so auditors tend to spend a considerable amount of time gaining assurance that the amount of the stated asset is reasonable

Return on Assets: $\frac{Profits\ after\ taxes}{Total\ assests}$

A measure of the return on total investment; It is sometimes desirable to add interest to after tax profits to form the numerator of the ratio since total assets are financed by creditors as well as by stockholders; hence, it is accurate to measure the productivity of assets by the returns provided to both classes of investors.

Current ratio: $\frac{Current\ assests}{Current\ liabilities}$

Indicates the extent to which the claims of short-term creditors are covered by assets that are expected to be converted to cash in a period roughly corresponding to the maturity of the liabilities.

Total assets turnover: $\frac{Sales}{Total\ assests}$

A measure of the utilization of all the firm's assets; a ratio below the industry average indicates the company is not generating a sufficient volume of business, given the size of its asset investment.

Inventory turnover: $\frac{Sales}{Investory\ of\ finished\ goods}$

When compared to industry averages, it provides an indication of whether a company has excessive or perhaps inadequate finished goods inventory.

Accounts Receivable (A/R) to sales ratio: Shows the relationship between unpaid sales and the total sales revenue. It is considered high if it is near to 1.0, because that means a significant amount of cash is tied up with the slow paying customers. Formula: Total accounts receivable (outstanding in an accounting period) ÷ sales revenue (in the same period).
Accounts Receivable (A/R) turnover = Average credit sale/Accounts Receivable: Ratio that shows the relationship between unpaid credit sales to total credit sales. It indicates, in general, the effectiveness (or lack of it) of a firm's credit policies and cash collection efforts. Formula: Outstanding accounts receivable (in an accounting period) ÷ credit sales revenue (in the same period), also called receivable turnover.

FINANCIAL RATIO ANALYSIS

Financial analysis techniques can help investigators discover and examine unexpected relationships in financial information (Davis, et al., 2010). These analytical procedures are based on the premise that relatively stable relationships exist among economic events in the absence of conditions to the contrary. Known contrary conditions which cause unstable relationships to exist might include unusual or nonrecurring transactions or events, and accounting, environmental, or technological changes. Public companies experiencing these events must disclose and explain the facts in their financial statements. Increasingly, private and not-for-profit companies follow best practices and do the same. Financial ratios are a great way to analyse a company's strengths and weaknesses. Ratios convert financial information to standardised format that can be used to compare with other companies and industry expectations. Unexpected deviations in relationships most likely indicate errors, but also might indicate illegal acts or fraud (Kovalerchuk et al., 2007). Therefore, deviations in

expected relationships warrant further investigation to determine the exact cause. Several methods of analysis assist the reader of financial reports in highlighting the areas that most likely represent fraudulent accounting methods (Dhar et al., 2010). An understanding of general relationships between certain financial statement balances is necessary to identify relationships that appear unusual. If sales increase, how should the cost of sales respond? If commission expense decreases, what would be expected of sales? Answers to questions such as these are the foundation of financial analysis. The following relationships are general, and traditionally occur between financial accounts; however, unique circumstances may render different results. The following is a worked example:

For the current year, ABC Ltd reported the following ratios:
-ROA = 20%
- Asset Turnover (being Sales/Average Total Assets) = 2.5 times
- Net Profit Margin = 15%

QUESTION: Why would we suspect at least one of these ratios is not correct?

We must be able to identify a relationship between the given ratios:
ROA = Net Income/Average Total Assets
ASSET TURNOVER = Sales/Average Total Assets
NET PROFIT MARGIN = Net Income/ Sales

Now, we can look at the common factors:
ROA and ASSET TURNOVER both use Average Total Assets

ROA and NET PROFIT MARGIN both use Net Income

ASSET TURNOVER and NET PROFIT MARGIN both use Sales

Now, we can look at the values provided:

ROA = 0.20 ASSET TURNOVER = 2.5 NPM = 0.15

We can then prove these values are not consistent. This can be done either by using algebra, or by simply assuming a value for one of the common factors. In this case, let's assume average total assets equal $100 (although any value can be chosen to prove if the ratios are consistent).

If that is the case; Net income/100 = 0.20 (so Net Income = 20)

Sale / 100 = 2.50 (so Sales = 250)

Then: Net Income / Sales would have to be: 20 / 250 =0.08

Conclusion: NPM of 15% is not consistent with the values of ROA = 20% and

Asset Turnover = 2.5 times.

These ratios may also reveal frauds other than accounting frauds. If an employee is embezzling from the company's accounts, for instance, the amount of cash will decrease disproportionately and the current ratio will decline. Liability concealment will cause a more favourable ratio. Similarly, a cheque-tampering scheme will usually result in a decrease in current assets, namely cash, which will, in turn, decrease the current ratio. In fact, these frauds may be more easily detected with ratio analysis because employees other than management would not have access to accounting cover-ups of non-accounting frauds. Anomalies in ratios could point directly to the existence of fraudulent actions. Accounting frauds can be much more subtle and demand extensive investigation beyond the signal that something is out of the norm.

THE PROPOSED FRAUD DETECTION MODEL

The Euclidean Distance based similarity metric is to evaluate the similarity of the standard industry ratios (*Return on Assets (ROA), Accounts Receivable (A/R) to sales ratio, Current Ratio, Total Asset Turnover and Inventory Turnover*) against the ratios of the current financial statement in order to determine if a fraud has occurred. Bajcsy and Kovačič (1989) argued that defining the problem will be the best way to understand the nature of the problem in its entirety. To evaluate the similarity between two different objects, *x* and *y*, a distance metric known as Euclidean Distance (*EU*) is used, this defines as follows: $EU\ (x,y) = \sqrt{(x - y)^2}$ (1)

This metric can be generalized into n-dimensions points, such that a={x_1, x_2, … x_n} and b={y_1, y_2, …, y_n}. In this case, n-dimensions *EU* metric is defined as: $EU\ (a,b) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$

$$= \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2} \qquad (2)$$

Let L_1 and L_2 be the existing standard industry ratios and the current financial statement ratios, respectively. Let x_i represent each ratio from the standard industry ratios and y_i represent each ratios from the financial statement, where $i=\{1, 2, \ldots, n\}$ and n is the total number of ratios. In this case, L_1 ={x_1, x_2, …, x_n} and L_2 ={y_1, y_2, …, y_n}. Euclidean distance can be normalized into a distance based similarity as follow: $S = \frac{1}{1+EU(L_1,L_2)}$ (3)

Similarity normalized *EU* into a value in between 0 and 1, where a value of 1 means that the two objects are identical and a value of 0 means that the two objects are not identical. This study focuses on detecting financial fraud and identifying the means of the fraud. In order to detect the fraud, the similarity between the industry standard ratios and the ratios of the current financial report is ranked. In doing so, the Euclidean distance between L_1 and L_2 is calculated first by using equation (1) and then the similarity can be ranked based on equation (3).

For example:
Using equation 1, if the industry standard for **Return on assets**: profit after tax/total assets = 1.5 and in the current financial report its 2.5 therefore it is calculated as follows:

$$EU\ (x,y) = \sqrt{(1.5 - 2.5)^2}$$
$$= \sqrt{1}$$
$$= 1$$

In order to normalise this into a value in between 0 to 1, use the similarity metric as showed in equation 3

$S = \frac{1}{1+EU(L_1,L_2)}$ = 0.5

The result is interpreted from predefined tabulated charts or user constructed tables based on empirical sample data constructed from use cases. The default setting is a result that falls between 0.5 and 1; and it indicates acceptance, but if it falls below 0.5 then it signals a red flag. A red flag means that the data requires further investigation. The metric acts as an indicator that eliminates possibilities rather than a deterministic measure that nominates an outcome. It cannot be used in isolation from expert knowledge and practitioner experience but in large datasets it can heighten the awareness of variations that are most likely indicate errors, and illegal acts or fraud. Therefore, deviations in expected relationships warrant further investigation to determine the exact cause.

CONCLUSION

Current digital forensics tools when applied to forensic accounting came up inadequate in our testing because they lack models that can sufficiently abstract from the data concerned. In this paper, the distance based similarity metric was proposed as the method for detection and identification of fraud, and a solution to the complexity problem. The result from the simple case developed to validate the method showed that the distance based similarity metric can detect financial statement variations that are an advancement on the simple ratio model tests and these variations can be mapped onto patterns of different activities, including fraud. It is also evident in this study that the metric effectively improved the performance, effectiveness and efficiency of the examination and analysis of large datasets of financial statements. This is helpful given the big data issues surrounding accounting and audit practice. For future work, a use case database of reference tables is to be developed from the application of this tool in industry so that the normalisation measure may be better matched onto scenario and context based situations. In its current form the metric is an improvement on the current ratio detection systems and has potential for coding into other digital forensic tools or into its own accounting forensic tool.

REFERENCES

Albano, P., Castiglione, A., Cattaneo, G., & De Santis, A. (2011). A Novel Anti-Forensics Technique for the Android OS. *Proceedings of the 2011 International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA)* (pp. 380-385). Maui: IEEE.

Albrecht, C. C. (2008). Fraud and Forensic Accounting In a Digital Environment. *White Paper for The Institute for Fraud Prevention, 1*(1), 1-32.

Ayers, R. (2007). *Cell phone forensic tools: An overview and analysis update*: Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology.

Baron, L. (2006). CPAs are a hot commodity. *Journal of Accountancy, 201*(2), 16.

Basuhail, A. A. S. (2010). Microsoft Excel as a tool for digital forensic accounting. *Proceedings of the 2010 International Conference on Information Retrieval & Knowledge Management* (pp.97-101). Malaysia: IEEE.

BusinessDictionary.com. (2016). *What is accounting?* Retrieved September 1, 2016, from http://www.businessdictionary.com/definition/accounting.html

Cohen, M. M., Crain, M. A., & Sanders, A. (1996). Skills used in litigation services. *Journal of Accountancy, 182*(3), 101.

Council, F. R. (2013). International Standard on Auditing (UK and Ireland) 610: Using the work of internal auditors. *The Financial Reporting Council Limited, 1*(1), 1-42.

Crumbley, D. L., Heitger, L. E., & Smith, G. S. (2005). *Forensic and investigative accounting* (Vol. 4025): CCH Incorporated.

Davis, C., Farrell, R., & Ogilby, S. (2010). Characteristics and skills of the Forensic Accountant. *American Institute of Certified Public Accountants*.

Dezfouli, F. N., Dehghantanha, A., Mahmoud, R., Sani, N. F. B. M., & bin Shamsuddin, S. (2012). Volatile memory acquisition using backup for forensic investigation. *Proceedings of the 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec)* (pp. 186-189). Kuala Lumpur: IEEE.

Dhar, P., & Sarkar, A. (2010). Forensic Accounting: An Accountant's Vision. *Vidyasagar University Journal of Commerce, 15*(1), 93-104.

DiGabriele, J. A. (2012). A Case Study on the Determination of Lost Profits for the Forensic Accountant. *Issues in Accounting Education, 27*(3), 751-759.

Glisson, W. B., Storer, T., & Buchanan-Wollaston, J. (2013). An empirical comparison of data recovered from mobile forensic toolkits. *Digital Investigation, 10*(1), 44-55.

Grispos, G., Storer, T., & Glisson, W. B. (2011). A comparison of forensic evidence recovery techniques for a windows mobile smart phone. *Digital Investigation, 8*(1), 23-36.

Guo, Y., & Slay, J. (2010). Data Recovery Function Testing for Digital Forensic Tools. In K.-P. Chow & S. Shenoi (Eds.), *Advances in Digital Forensics. Sixth IFIP WG 11.9 International Conference on Digital Forensics, Revised Selected Papers* (pp. 297-311). Heidelberg: Springer

Haber, J. R. (2004). CHAPTER 2: Financial Statements. In, *Accounting Demystified* (pp. 4-12). American Management Association International.

Houck, M. M., Kranacher, M.-J., Morris, B., & Riley Jr, R. A. (2006). Forensic accounting as an investigative tool. *The CPA Journal, 76*(8), 68.

Jansen, W., & Ayers, R. (2007). Guidelines on cell phone forensics. *NIST Special Publication, 800*, 101.

Kovalerchuk, B., Vityaev, E., & Holtfreter, R. (2007). Correlation of complex evidence in forensic accounting using data mining. *Journal of Forensic Accounting, 8*(1).

Kubi, A. K., Saleem, S., & Popov, O. (2011). Evaluation of some tools for extracting e-evidence from mobile devices. *Proceedings of the 2011 5th International Conference on Application of Information and Communication Technologies (AICT)* (pp.1-6). Baku: IEEE.

Mohtasebi, S., & Dehghantanha, A. (2013). Towards a Unified Forensic Investigation Framework of Smartphones. *International Journal of Computer Theory and Engineering, 5*(2), 351-355.

Morrissey, S. (2010). iOS Operating and File System Analysis. In *iOS Forensic Analysis for iPhone, iPad, and iPod touch* (pp. 25-66). NY: Apress.

NIST. (2013). *Test Results for Mobile Device Acquisition Tool: Device Seizure v5.0 build 4582.15907*. Retrieved from: https://www.ncjrs.gov/pdffiles1/nij/241153.pdf

Nissan, E. (2012). The Forensic Disciplines: Some Areas of Actual or Potential Application [Nissan2012]. In *Computer Applications for Handling Legal Evidence, Police Investigation and Case Argumentation* (pp. 841-989). Dordrecht: Springer.

NIST. (2001). General Test Methodology for Computer Forensic Tools. *NIST Technical Report Ver1.9, 1*(1), 1-8.

Panigrahi, P. K. (2006*). Discovering fraud in forensic accounting using data mining techniques*. Chartered Accountant: New York.

Reinstein, A., & McMillan, J. J. (2004). The Enron debacle: more than a perfect storm. *Critical Perspectives on Accounting*, *15*(6–7), 955-970.

Rezaee, Z., & Burton, E. J. (1997). Forensic accounting education: insights from academicians and certified fraud examiner practitioners. *Managerial Auditing Journal*, *12*(9), 479-489.

Smith, M., Sagafi-Nejad, T., & Wang, K. (2008). Going international: Accounting and auditing standards. *Internal Auditing*, *23*(4), 3-12.

Wells, J. T. (2003). The fraud examiners. *Journal of Accountancy*, *196*(4), 76.

Willis, V. F. (2016). A model for teaching technology: Using Excel in an accounting information systems course. *Journal of Accounting Education*, *36*, 87-99.

Yannikos, Y., Franke, F., Winter, C., & Schneider, M. (2011). 3LSPG: Forensic Tool Evaluation by Three Layer Stochastic Process-Based Generation of Data. In H. Sako, K. Y. Franke, & S. Saitoh (Eds.), *Computational Forensics: 4th International Workshop, IWCF 2010, Tokyo, Japan, November 11-12, 2010, Revised Selected Papers* (pp.200-211). Heidelberg: Springer.

GOOGLE EARTH FORENSICS ON IOS 10'S LOCATION SERVICE

Brian Cusack[1], Raymond Lutui[2]
[1]Auckland University of Technology, Auckland New Zealand, [2]Christ's University in Pacific, Tonga
brian.cusack@aut.ac.nz, raymond.lutui@aut.ac.nz

Abstract
The easy access and common usage of GNSS systems has provided a wealth of evidential information that may be accessed by a digital forensic investigator. Google Earth is commonly used on all manner of devices for geolocation services and consequently has a wide range of tools that will relate real time and stored GNSS data to maps. As an aid to investigation Google Earth forensics is available for use. An investigator can use it by downloading geolocation data from devices and placing it on Google Earth maps, place geolocation data on historical archival maps, or by direct usage of the application in a device. In this paper we review the Google Earth forensics tool and use a simplistic scenario to demonstrate the power of the application for courtroom walk-throughs. The entry-level tool is free and can be used effectively to enhance the presentation of geolocation data.

Keywords
Digital Forensics, Location based service, GNSS, Google Earth forensics, Investigation

INTRODUCTION

Advancements in technology over the last 20 years have drastically altered the way people live and do business (Garfinkel, et al., 2009, p.3). The advancements have driven the growth of popularity and usages of mobile devices such as Smartphones and all the related application packages. This growth is expected to continue for the foreseeable future and the mobile devices market to expand (Liu, et al., 2012, p.145). The increasing storage capacity and functionality of these mobile devices have outnumbered personal computers and become the personal device of choice in our society. Large amounts of information from phone book to photo albums and videos, to emails and text messages, to financial records and GNSS (Global Satellite Positioning System) records, are stored in these phones (Mylonas, et al., 2012, p.249). Mobile devices have the ability to network especially via wireless connections such as Wi-Fi, cellular network such as 3G and 4G (Liu, et al., 2012). Mobile devices are relatively small, portable and widely used by all ages, and especially by young people. Mobile devices consist of Smartphones, Tablets, embedded devices, Personal Digital Assistants (*PDA*), and a range of evolving products (Bennett, 2012, p.159). The personalisation and control handed to the user make these anytime, anywhere devices the communication device of choice. The end user finds the usefulness, capability, numerous applications, high processing power and improved communication performance, to be in line with their demand for experience (Yusoff, et al., 2014, p.141).

Mobile phone forensics is defined by the National Institutes of Standards and Technology as, the science of recovering digital evidence from a mobile phone under forensically sound conditions using accepted methods (Jansen & Ayers, 2007, p.6). The new generation of digital mobile devices are known as Smart devices because of their processing power, memory and storage spaces that are very similar to that of a desktop computer. However, the operating system that runs on these Smart devices has its own file system and structure. On Smart devices, the operating system functions come from volatile memory whereas with computers it is in RAM and ROM. These Smart devices are capable of storing, transmitting and processing large amounts of private and confidential data (Owen & Thomas, 2011, p.25). Yet, the main challenge to forensic experts is the fact that Smart devices are different proprietary designs, software, and access controls. As a result, forensic experts must stay up-to-date with the latest technology in order to be able to adapt to technological changes and strategize new approaches (Rajendran and Gopalan, 2016, p.394).

The affordability of these devices, the proliferation of wireless hotspots and the availability of wireless location services have created new business opportunities and demand for use (Karygiannis and Antonakakis, 2009, p.308). This has led the researchers in the field to find a way of providing location based services to mobile clients.

This paper is structured to review previous literature on geolocation services, to review Google Earth forensic capability, and then to provide a test scenario and walk-through. The results are then discussed to elicit the

implications for practice. The conclusion is that Google Earth is a powerful tool to visualise location data in context. However, it is important to practice with the tool and then to use forensic examination procedures to collect and to manage the evidence. This ensures the best opportunity that the findings can be admissible in the court of law.

PREVIOUS LITERATURE

Recently, the mobile sensing data are used for analysing users' activities such as, their emotions, health conditions, their usage patterns and social relationships. This is done in order to study and understand human's behaviour (Mo, et al., 2015, p.391). These studies were triggered by the widespread use of highly capable and multifunctional Smart devices, collecting and analysing mobile sensing data are given more attention in recent years. These days, Smartphones are very much part of people's everyday lives and including recreation. User's takes these Smart devices along everywhere they go. These Smart devices such as Smartphones consist of various sensors and technologies. As a result, they collect a variety of information such as the user's locations, activities, phone calls, SMS, e-mails, call logs, contact list, applications, and so on. Recently, various smart application services have been developed using GNSS, RFID (Radio Frequency Identification) and sensor network connectivity.

The GNSS has been successfully applied for outdoor location tracking by many applications, but it might still be insufficient in an indoor environment where GNSS signals are often severely obstructed.

Location Based Services

Researchers in the field have been working on the possibility of providing location based services (LBSs) to mobile clients. Researchers believed that such technology can be very helpful if available indoors in places like shopping malls, train stations, airports and universities for sharing information, schedule change of train and flight, and so on (Sadhukhan, et al., 2010, p.10). As a result, the indoor location tracking systems has shown its importance for personal information use. GNSS is the well-known and commonly used technology for providing location information services. Location Based Services (LBS) on the other hand, has limitations of dependency on satellite visibility and problems with accuracy and consistency of service in cities, covered areas and multi-path contexts (Kim, et al., 2014, p.89). Mobile RFID applications has been utilised to play a role in the implementation of an indoor tracking system for exhibition service (Kim et al., 2014, p.96). RFID-Based Navigation System is utilised to reduce localization errors (Reza and Buehrer, 2012, p.1023). However, the intensity of experience that satisfies end user geolocation requirements is driving use of positioning technology for security applications, quality control in manufacturing, and safety applications in high-risk areas (Werner, 2014a, p.74).

Location-based services are built by detecting the environments of a mobile device. The inherent complexities of buildings and the localization problem inside buildings that make most indoor location based services use a large amount of environmental information from various sources (Werner, 2014b, p.169). Nonetheless, it's always being sufficient to detect the location of the user with regards to the planned route. The navigational events consist of directional changes yet, inside buildings, the people are missing the navigational experience (Werner, 2014b, p.169).

Mobile Sensing

In order to understand human's behaviours, mobile sensing data is collected and analysed to show activities, usage patterns, emotions, health conditions and social relationships (Mo, et al., 2015, p.391). The customer location tracking system is also used to send location specific information advertisements to users via their mobile Smart device (Keikhosrokiani, et al., 2011, p.527). A mobile smart device such as Smartphones has the characteristics of personalization. As a result, they are utilised to develop a platform for home care for the elderly combining with alarm clock. This allows the device to remind the patients to take their medicine (Cheng, et al., 2011, p.259). Sensor-equipped mobile device revolutionised sectors of our economy including business, healthcare, social networks, environmental monitoring, and transportation (Lane et al., 2010, p.140). Also Ultra-High Frequency (UHF) RFID technology has been applied to supply chain, asset tracking, antifraud system and intelligent transportation systems and so on (Eo, et al., 2008, p.730). RFID technology on Smart devices incorporates the RFID technology with the cellular network. This allows access to information stored on the tag through the RFID reader on the Smart device such as Smartphone (Peng, et al., 2012, p.243). RFID technologies have been utilised in retail to track inventory, in manufacturing to track product status, and also in airlines to track lost baggage (Wang, et al., 2009, p.495). However, researchers in the field have identified two primary types of GNSS location traces - location-only and location-based. Location uses the reference data for visual

inspection to infer trip purposes (Wolf et al., 2003, p.6). Location-based studies on the other hand, incorporate other supplementary data such as related trip information (Schönfelder et al., 2003, p.8; Stopher et al., 2008, p.2).

Location Value

Location plays a key role in defining the nature of human activities. Location can determine users' requirements, buying behaviours and service choices (Rao and Minakakis, 2003, p.63). If the provider knows the user's exact location and has the ability to target valuable information, the benefits can be reciprocated. However, awareness of the user's location is only part of the problem (Schönfelder, et al., 2003, p.1). The Human trajectories3 application has sets of time-stamped locations describing individuals' movements in time (typically over a day), and are potentially of great use for many different interested parties. On their own, they reveal the fundamentals of human mobility patterns (Gonzalez et al., 2008, p.780). Therefore, human-carried mobile devices, routing and broadcasting algorithms deployed in various cellular networks should take advantage of the properties of human mobility in order to be effective (Marin et al., 2014, p.204). Thus, based on the where, how, why, when, and who with, users' are routing. Someone can be in a rugby game with friends most likely to have different LBS profile from that of a person on a routine trip to the supermarket or shopping mall. If the provider has the ability to deliver reliable information to users, they might have difficulty identifying what the user is doing or looking for at that location (Heinemann and Gaiser, 2015, p.55). All of these matters can become evidential and in this day and age they are stored digitally.

Significant developments and advances in the field of wireless communication technologies, location sensors and global networking, are driving a new world (Saha and Mukherjee, 2003, p.25). The Smartphone generation is accustomed to the connected world; it becomes a custom to be connected to any person, anytime, anywhere. In fact, people want to be able to monitor and control everything in their lives at anytime from anywhere (Siewiorek, 2012, p.323). However, in the inter-connected world, people prefer not to only limit to Smart devices only. The Internet is being extended to connect every device around us, and is generally known as the Internet of Things (IoT) and the Internet of Everything (IoE). IoT is a concept of a dynamic that is constantly building up and escalating into a future technology full immersion experience (Jun, et al., 2011, p.3). According to Kantarci and Mouftah (2015), IoT interconnects billions of objects that are uniquely identifiable and that have communications, computing and sensing functionalities (p.1865). In such an environment, the acquired data mainly used to analyse human's behaviour, looking for patterns that can be useful in location based service (Vlassenroot, et al., 2015, p.19). However, in some cases even if the data is stored offline, the size of the data is so large and distributed; it will require the use of big data analytical tools for processing (Aggarwal et al., 2013, p.383). Big data comes from various sources, in this case, the location based service sensor, the mobile device GNSS signals, sensors used to gather climate information, social networking sites data, digital photos and videos, and so on. These data is known as *Big Data*, and analysing such data sets strengthens new streams of productivity growth, innovation, and consumer surplus. Learning gained from the result of analysing *Big Data* can help to answer human technology immersion and proximity questions (Ciobanu, et al., 2014, p.4). In addition all of this data and information can become evidential and useful to forensic investigation.

GOOGLE EARTH REVIEW

Google Earth Forensics is a practical and easily accessible tool kit for using with Google Earth Geo-Location services. Google Earth (GE) is a tool that enables its users to view the planet through a virtual globe. Users can navigate through satellite images, aerial photography, and even views of street level imagery and 3D models of the world. This includes locations like oceans, the moon, Mars, and outer space. Features in GE allows tours of locations, historical archive mapping, and to fly across locations using a flight simulator (Harrington and Cross, 2015, p.1). GE provides a representation to real locations within the material limitations of the data collected and the visual presentation. When a location is entered into GE, a map will display and it will include the labelled position of the place you are searching for. It will also allow zooming to see 3D structures or actual photos of a location. GE also allows the viewing of areas of the earth using custom maps or overlays which contain data imported from GNSS units and other digital devices. It is an information source and geolocation search application for multiple usages (Parks, 2009, p.537).

GE is known to be utilised by Police throughout the world in various ways to investigate crimes and share information with the general public. Analysts in law enforcement agencies gather data from police reports and other sources and made it available through GE. For instance, the Shawnee Police Department in Kansas provides data that can be loaded into GE to see locations where robberies, auto thefts, vandalism and other crimes have taken place (Shawnee PD, 2010, p.1). In the digital forensics domain, knowing where a computer or device has been is very important in an investigation. However, we are living in an age where mobility is highly vital. As a result, knowing the exact location of a device when involved in a certain activity is vital information

to a forensic investigator (Google Earth Forensics, 2015, p.1). It is possible to chart user's movements from data found on Smartphones from metadata of a photo and other media. The data by itself, however, may not mean much. It needs to be translated into a meaningful form. Perkins and Dodge (2009) argued that, secrets are strongly associated with visual culture, where they are hidden from view but may be revealed. They are ubiquitous, but often unseen and are particularly associated with certain spaces. Google Earth Forensics (2015) note that, many devices store location data as a matter of course – even, sometimes, when the user has asked it not to (p.546).

Forensics is the use of scientific or technological techniques to investigate and establish facts. In a criminal case, the facts you are looking for will be evidence of how a crime was committed and who was responsible (Harrington and Cross, 2015, p.4). Throughout a process of preserving the crime scene and identifying, gathering and examining evidence, information is carefully documented. This is used in the hopes of understanding what occurred, and so that it may be used to identify, arrest and convict the perpetrator (Takeuchi, et al., 2012, p.183). GE can be used to search and display location-specific information in a way that is more telling than the raw data. Using GE as a forensic tool, you can also:

- Import data from mobile devices and GNSS units to determine a route that was taken, or locations that a person visited
- Determine the location where a photo was taken using geo-location information stored in a digital picture
- Create maps that display locations that a device visited, and movies that covey location-based information in a compelling format for investigations and court presentation

The following screenshots illustrate the use of Google forensics for the processing of data taken from a mobile device. XRY (figure 1) is used as the extraction tool and then the data is imported into Google forensics to demonstrate its usefulness. The screen shot show a step-by-step process guide for the use of the tool and instruction for others who wish to use it.
Harrington and Cross (2015) stated that, GE can be incorporated into various phases of the digital forensic procedure. Most often, users may surprise to know that it is used in later parts of a case, when coordinates from various sources needs to be analysed or a tool to create presentations relating to geographic locations. In some cases, it may also be used to acquire GNSS data from a device, although other tools may be more suited to collecting such data for a forensic investigation (p.73).

In this example, the XRY was used to acquire data from an iPhone 6 plus 64GB running on iOS 10.0.2. Figure 1 showed XRY logical was used to acquire the data from the iPhone 6 Plus. The XRY Reader v6.18 - 32bit as showing in figure 2 is used to read and analyse the data. The forensic process was done on HP laptop running on i7 processor with 8GB of RAM. Operating system is Microsoft Windows 10 Professional.

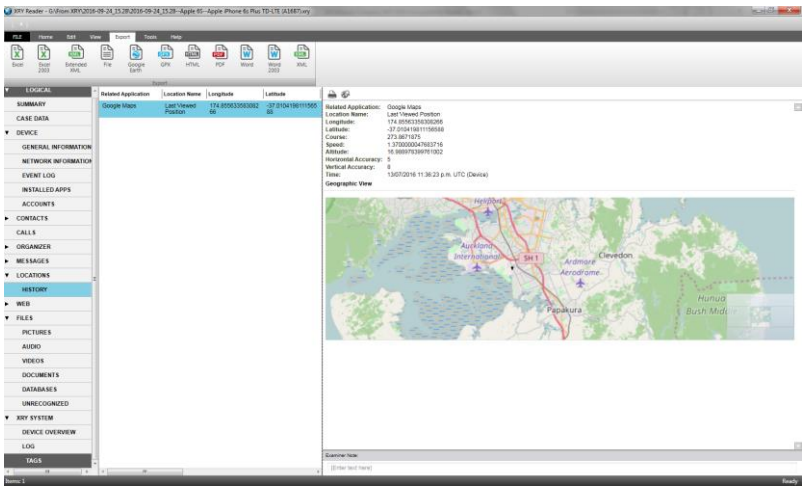


Figure 1: Preliminary view provided by XRY

The XRY Reader in figure 1 showed detailed information on the location data found and selected in the data acquired. This information includes longitude, latitude, altitude, and also date, time, and a preview of the location in google map. This location data can also be exported in various formats such as pdf, word, excel, GPX, XML, HTML or GE.

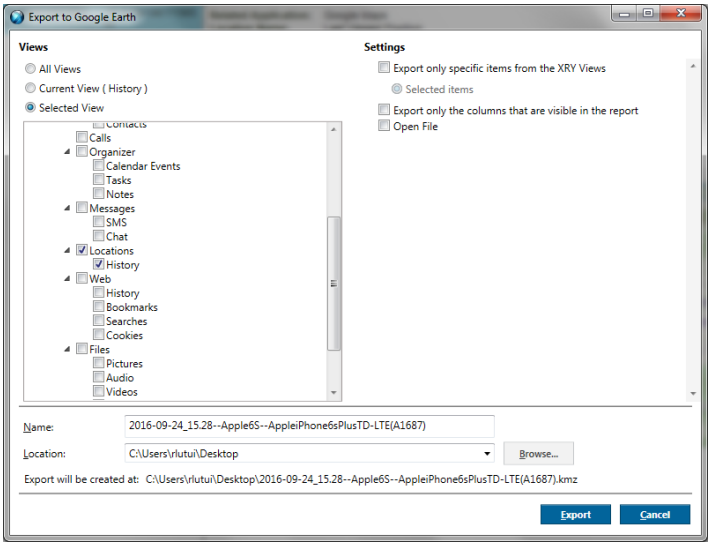


Figure 2: Exporting location data into Google Earth format

In this study, the location data was exported into GE format as shows in figure 2 and then opened with GE. For this study, GE 7.1.7.2600 was used.

Figure 3 shows the result when the exported file is opened in GE. This shows a much clearer view of the location showed in figure 2. GE itself is not a recognized or standardized forensic tool but has the functionality to display geolocation data. All of the issues with accuracy and timeliness require declaration and any variations reported. Figure 4 has the Google forensics import GNSS data panel.

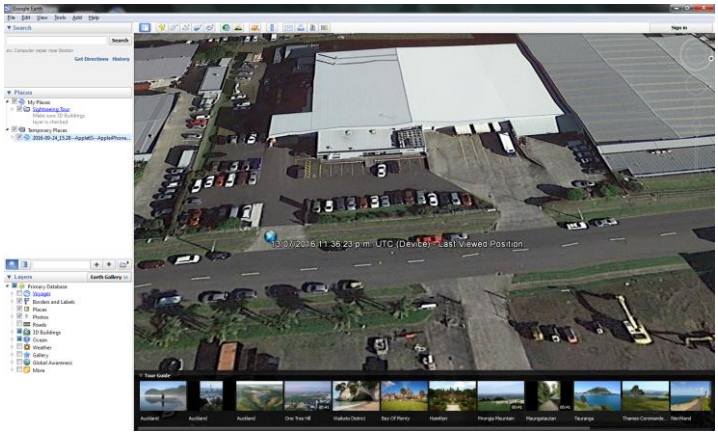


Figure 3: Google Earth view

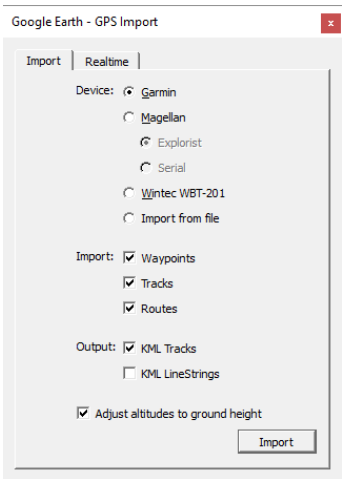


Figure 4: Google Earth GNSS import panel

Compliance is important because any matter may be cross-examined or challenged in court. For instance, instead of simply plugging the GNSS device directly into a USB port, it is highly recommended to make sure that a software write protection or a hardware write blocker is used to prevent any accidental modification of data. Acquiring the GNSS location data in this manner, it shows the number of waypoints, tracks and routes that are imported from a GNSS device. However, importing GNSS data in this way has disadvantages because it copies the data directly off the device into GE. It only performs a logical acquisition and does not retrieve any deleted or even hidden data on the device.

To import the data direct in to GE, perform the following steps.

1) In GE, click on the *Tools* menu, and then click the *GNSS* menu item
2) When the *GNSS Import* dialog box appears select the type of device you are importing from. As our scenario involves a Garmin GNSS unit, we would click the *Garmin* option to import from that device
3) In the *Import* section, select what you want to import: n Waypoints – these are individual locations marked on the GNSS unit. n Tracks – these are where the GNSS has been. n Routes – these are a series of locations where the user wishes to navigate.
4) In the *Output* section, select the output to be used with GE.

When ready to start obtaining data, clicking the Import button, GE will import the data from the device. A summary of the imported data will be displayed on the screen once finished.

APPLICATION OF GOOGLE EARTH FORENSICS

In the previous section a step-by-step guide has been provided on how to use GE with Google forensics. In this section, a simple scenario is adopted to demonstrate how generated data can be displayed as a courtroom walk-through.

"Alleged young children kidnapping and trafficking"

A deal gone wrong and from the crime scene, a mobile Smart device was found and suspected of being involved in some criminal activities. As a result, the Smart phone was taken back to the lab for processing. Upon examination of the Smart phone, there were several geotagged photographs which led the investigator to look at the geolocation data on the device.

In this case, the geolocation file has been acquired from the device, and the location data can be imported into GE. In GE, the placemarks can also be modified. This is a commonly used feature and very important for pinpointing locations and presenting information regarding a crime scene in this case or even other places related to a case. To add a placemark in GE, click on the *Add Placemark* button on the toolbar.

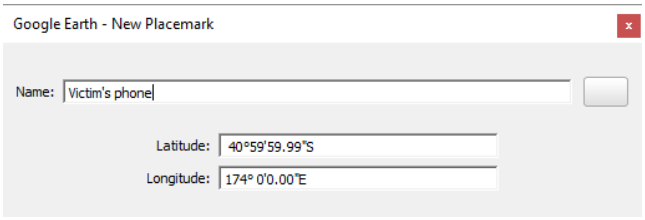


Figure 5: Customising placemark dialog

Fill out the upper portion of the dialog with the information shown in Figure 5 and click *OK* when done. Once the location file from the Smart device is loaded into GE, it displays the waypoints as shown in figure 6.

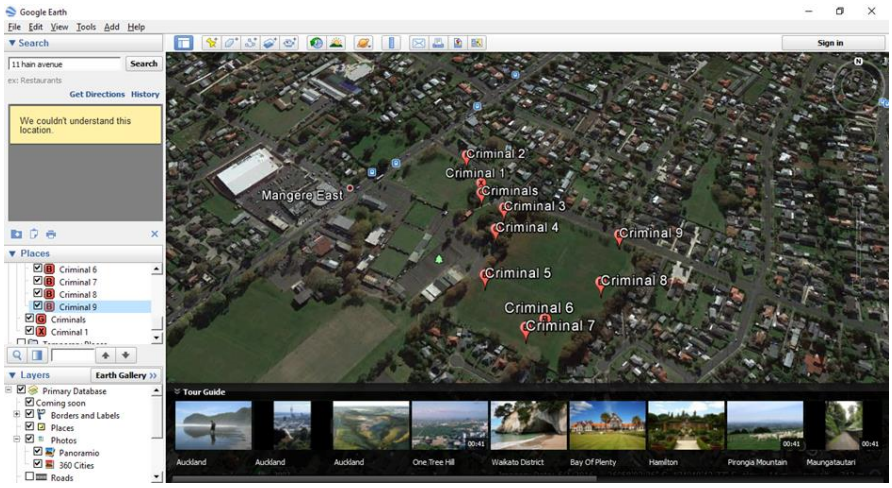


Figure 6: Waypoints

In relation to the case scenario, the illustration in figure 6 showed the places in the public park where the events are located. Zooming in to get a ground level view, allows a representation of the crime scene in three-dimensional space. The images can be overlaid with evidence as it has been collected and GNSS located. Placemarks in GE can be customised according to the nature of the case depending whether it is a phone, a vehicle or other matter at the crime scene. It is not only the placemarks but also content may be added to describe area around the placemark to link it to reports and photographs and narratives. In this way three-dimensional and historical image mapping may be used to reconstruct a crime scene and presented in visual format in a courtroom.

CONCLUSION

Access to Google forensics is extremely helpful to all forensic investigators. The entry-level packages is free and provides considerable capability to process and demonstrate geolocation data. Geolocation and psycho location data are readily available from Smart devices in this day and age and provide a whole multidimensional image of human behaviour. In this paper we have provided a step-by-step guide as to how to use the tool and to apply it to a courtroom walk-through. We have also cautioned that acceptable forensic practices must also be used for the data acquisition and management processes. Further research and demonstration of historical archiving Google forensics capability can also be helpful in cold cases.

REFERENCES

Aggarwal, C. C., Ashish, N., & Sheth, A. (2013). The Internet of Things: A Survey from the Data-Centric Perspective [Aggarwal2013]. In C. C. Aggarwal (Ed.), *Managing and Mining Sensor Data* (pp. 383-428). Boston: Springer.

Bennett, D. (2012). The challenges facing computer forensics investigators in obtaining information from mobile devices for use in criminal investigations. *Information Security Journal: A Global Perspective, 21*(3), 159-168

Cheng, R.-S., Ke, C.-Y., Tsai, C.-Y., & Wang, C.-J. (2011). A Mobile Homecare Application Combining with Alarm Clock and GNSS Positioning Function. In R.-S. Chang, T.-h. Kim, & S.-L. Peng (Eds.), *Proceedings of the SUComS 2011, Hualien, Taiwan Second International Conference on Security-Enriched Urban Computing and Smart Grid.* (pp.259-268). Heidelberg: Springer.

Ciobanu, R.-I., Cristea, V., Dobre, C., & Pop, F. (2014). Big Data Platforms for the Internet of Things. In N. Bessis & C. Dobre (Eds.), *Big Data and Internet of Things: A Roadmap for Smart Environments* (pp. 3-34). Cham: Springer.

Eo, Y., Bang, H., Choi, K., Jeon, S., Jung, S., Lee, D., & Lee, H. (2008). A Single-Chip CMOS Transceiver for UHF Mobile RFID Reader. *IEEE Journal of Solid-State Circuits, 43*(3), 729-738

Garfinkel, S., Farrell, P., Roussev, V., & Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. *Digital Investigation, 6, Supplement*, S2-S11.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *Nature, 453*(7196), 779-782.

Google Earth Forensics. (2015). *Network Security, 2015*(3), 4.

Harrington, M., & Cross, M. (2015). *Google Earth Forensics: Using Google Earth Geo-Location in Digital Forensic Investigations*. Waltham: Elsevier.

Heinemann, G., & Gaiser, C. (2015). Location-based services as Base Factor No. 2 for SoLoMo. In *Social - Local - Mobile: The Future of Location-based Services* (pp. 55-99). Heidelberg: Springer.

Jansen, W., & Ayers, R. (2007). Guidelines on cell phone forensics. *NIST Special Publication, 800*, 101.

Jun, Z., Simplot-Ryl, D., Bisdikian, C., & Mouftah, H. (2011). The internet of things. *IEEE Commun. Mag, 49*(11), 30-31.

Kantarci, B., & Mouftah, H. T. (2015). Sensing services in cloud-centric Internet of Things: A survey, taxonomy and challenges *Proceedings of the 2015 IEEE International Conference on Communication Workshop (ICCW)* (pp.1865-1870). London: IEEE.

Karygiannis, A., & Antonakakis, E. (2009). Security and Privacy Issues in Agent-Based Location-Aware Mobile Commerce. In M. Barley, H. Mouratidis, A. Unruh, D. Spears, P. Scerri, & F. Massacci (Eds.), *Safety and Security in Multiagent Systems: Research Results from 2004-2006* (pp. 308-329). Berlin: Springer.

Keikhosrokiani, P., Mustaffa, N., Sarwar, M. I., Kianpisheh, A., Damanhoori, F., & Zakaria, N. (2011). A Study towards Proposing GNSS-Based Mobile Advertisement Service. In A. Abd Manaf, A. Zeki, M. Zamani, S. Chuprat, & E. El-Qawasmeh (Eds.), *Proceedings of the ICIEIS 2011 International Conference on Informatics Engineering and Information Science: Part II, Kuala Lumpur, Malaysia* (pp. 527-544). Heidelberg: Springer.

Kim, S. H., Park, H., Bang, H. C., & Kim, D.-H. (2014). An indoor location tracking based on mobile RFID for smart exhibition service. *Journal of Computer Virology and Hacking Techniques, 10*(2), 89-96.

Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., & Campbell, A. T. (2010). A survey of mobile phone sensing. *IEEE Communications Magazine, 48*(9), 140-150.

Liu, H., Azadegan, S., Yu, W., Acharya, S., & Sistani, A. (2012). Are We Relying Too Much on Forensics Tools? In R. Lee (Ed.), *Software Engineering Research,Management and Applications 2011* (pp. 145-156). Berlin: Springer.

Marin, R.-C., Ciobanu, R.-I., Dobre, C., & Xhafa, F. (2014). Techniques and Applications to Analyze Mobility Data. In F. Xhafa & N. Bessis (Eds.), *Inter-cooperative Collective Intelligence: Techniques and Applications* (pp. 203-237). Heidelberg: Springer

Mo, X., Shi, D., Yang, R., Li, H., Tong, Z., & Wang, F. (2015). A Framework of Fine-Grained Mobile Sensing Data Collection and Behavior Analysis in an Energy-Configurable Way. *Proceedings of the 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)* (pp. 391-398). Chengdu: IEEE.

Mylonas, A., Meletiadis, V., Tsoumas, B., Mitrou, L., & Gritzalis, D. (2012). Smartphone Forensics: A Proactive Investigation Scheme for Evidence Acquisition. In D. Gritzalis, S. Furnell, & M. Theoharidou (Eds.), *Information Security and Privacy Research: Proceedings of the 27th IFIP TC 11 Information Security and Privacy Conference, SEC 2012, Heraklion, Crete, Greece,* (pp. 249-260). Berlin: Springer.

Owen, P., & Thomas, P. (2011). An analysis of digital forensic examinations: Mobile devices versus hard disk drives utilising ACPO & NIST guidelines. *Digital Investigation, 8*(2), 135-140.

Parks, L. (2009). Digging into Google Earth: An analysis of "Crisis in Darfur". *Geoforum, 40*(4), 535-545.

Peng, Q., Zhang, C., Song, Y., Wang, Z., & Wang, Z. (2012). A low-cost, low-power UHF RFID reader transceiver for mobile applications *Proceedings of the 2012 IEEE conference on Radio Frequency Integrated Circuits* (pp.243-246). Montréal: IEEE.

Perkins, C., & Dodge, M. (2009). Satellite imagery and the spectacle of secret spaces. *Geoforum, 40*(4), 546-560.

Rajendran, S., & Gopalan, N. P. (2016). Mobile Forensic Investigation (MFI) Life Cycle Process for Digital Data Discovery (DDD). In P. L. Suresh & K. B. Panigrahi (Eds.), *Proceedings of the International Conference on Soft Computing Systems: ICSCS 2015, Volume 2* (pp. 393-403). New Delhi: Springer.

Rao, B., & Minakakis, L. (2003). Evolution of mobile location-based services. *Commun. ACM, 46*(12), 61-65.

Reza, Z., & Buehrer, R. M. (2012). Autonomous Mobile Robot Navigation Systems Using RFID and their Applications. In *Handbook of Position Location:Theory, Practice and Advances* (pp. 1023-1054): Wiley-IEEE Press.

Sadhukhan, P., Sen, R., & Das, P. K. (2010). A Middleware Based Approach to Dynamically Deploy Location Based Services onto Heterogeneous Mobile Devices Using Bluetooth in Indoor Environment. In C.-C.

Chang, T. Vasilakos, P. Das, T.-h. Kim, B.-H. Kang, & M. Khurram Khan (Eds.), *Proceedings of the ACN 2010 Second International Conference on Advanced Communication and Networking* (pp. 9-22). Heidelberg: Springer.

Saha, D., & Mukherjee, A. (2003). Pervasive computing: a paradigm for the 21st century. *Computer, 36*(3), 25-31.

Schönfelder, S., Ethz, I., & Samaga, U. (2003). Where do you want to go today?–More observations on daily mobility*Citeseer*. Symposium conducted at the meeting of the *Proceedings of the 3 rd Swiss Transport Research Conference on Session Mobility*, Monte Verità: Citeseer.

Siewiorek, D. (2012). Generation smartphone. *IEEE Spectrum, 49*(9), 54-58.

Stopher, P., Clifford, E., Zhang, J., & FitzGerald, C. (2008). Deducing mode and purpose from GNSS data. *Institute of Transport and Logistics Studies*, 1-13.

Shawnee PD. (2010). *Crimes maps: Google Earth (kmz) 2010 crime reports*. Retrieved September 30, 2016, from http://www.cityofshawnee.org/WEB/PoliceCMS.nsf/c0019294e957d2c28525754a004b58b4/4b7c35995b121854862575e5004a6574?OpenDocument

Takeuchi, T., Matsuki, R., & Nashimoto, M. (2012). GNSS cell phone tracking in the Greater Tokyo Area: A field test on raccoon dogs. *Urban Ecosystems, 15*(1), 181-193.

Vlassenroot, S., Gillis, D., Bellens, R., & Gautama, S. (2015). The Use of Smartphone Applications in the Collection of Travel Behaviour Data. *International Journal of Intelligent Transportation Systems Research, 13*(1), 17-27.

Wang, J., Zhang, C., Chi, B., Wang, Z., & Wang, Z. (2009). A fully integrated CMOS UHF RFID reader transceiver for handheld applications *Proceedings of the 2009 IEEE Conference on Custom Integrated Circuits* (pp.495-498). CA: IEEE.

Werner, M. (2014a). Basic Positioning Techniques. In *Indoor Location-Based Services: Prerequisites and Foundations* (pp. 73-99). Cham: Springer.

Werner, M. (2014b). Event Detection for Indoor LBS. In *Indoor Location-Based Services: Prerequisites and Foundations* (pp. 169-179). Cham: Springer.

Wolf, J., Oliveira, M., & Thompson, M. (2003). Impact of underreporting on mileage and travel time estimates: Results from global positioning system-enhanced household travel survey. *Transportation Research Record: Journal of the Transportation Research Board* (1854), 189-198.

Yusoff, M. N., Mahmod, R., Dehghantanha, A., & Abdullah, M. T. (2014). Advances of Mobile Forensic Procedures in Firefox OS. *International Journal of Cyber-Security and Digital Forensics (IJCSDF), 3*(4), 183-199.

A FORENSIC EXAMINATION OF SEVERAL MOBILE DEVICE FARADAY BAGS & MATERIALS TO TEST THEIR EFFECTIVENESS

Ashleigh Lennox-Steele, Alastair Nisbet
Security & Forensic Research Group, Auckland University of Technology
Auckland, New Zealand
a.lennoxsteele@gmail.com, alastair.nisbet@aut.ac.nz

Abstract

A Faraday bag is designed to shield a mobile phone or small digital device from radio waves entering the bag and reaching the device, or to stop radio waves escaping through the bag from the device. The effectiveness of these shields is vital for security professionals and forensic investigators who seize devices and wish to ensure that their contents are not read, modified or deleted prior to a forensic examination. This research tests the effectiveness of several readily available Faraday bags. The Faraday bags tested are all available through online means and promise complete blocking of all signals through the bag. Additionally, other materials that can be used if a Faraday bag is not available, such as tin foil and a tin can are tested and compared with the Faraday bags. A selection of common mobile phones from various manufacturers is tested in the shielding material. Additionally, 3G / 4G, WiFi and Bluetooth are tested with the bags and materials on those so equipped devices to ascertain whether the material blocks all signals from communicating technologies on the phones. Results show that performance of the bags is not as promised by most vendors and that in urgent situations other materials at hand may suffice to perform the same function as a Faraday bag.

Keywords
privacy, security, shielding, mobile, forensics, Faraday bags

INTRODUCTION

The utilisation of smart mobile devices plays a key role in the majority of day-to-day lives, where most carry at least one electronic device, whether they be for corporate use as a portable office, or personal use for tasks such as social networking and entertainment (Rajendran & Gopalan, 2016). This reliance on smartphones is credited to their proven advantages regarding efficiency in fulfilling both personal needs and business development, supported by the influx of availability and variety of mobile devices and their compatible applications (Khan, Abbas, & Al-Muhtadi, 2015). Current mobile devices, specifically the smartphone, combine telephony and computer services in a convenient handheld device. Standard telephone services are provided through a cellular network, while Internet services are utilised through a Wi-Fi connection or via 3G and 4G cellular data networks (Soukup, 2015). Google's Android, and Apple's iOS are the two main platforms which dominate today's mobile market (FireEye, 2015), with the everyday use of smart devices steadily increasing, while the utilisation and cost of basic feature phones and personal computers continually fall (Ophoff & Robinson, 2014). Where older models of mobile phones could only store a limited amount of data, which was easily obtainable by forensic investigators, the increased popularity and development of smart devices has complicated the techniques which must be employed by forensic investigators in order to retrieve the many variants of stored data in a forensically sound manner (Bennett, 2012).

LITERATURE REVIEW

A mobile device must have at least one wireless network interface to allow for data communications, whether it be Wi-Fi, cellular networking or additional technologies which can be used for connecting a device to networks or the Internet, as well as providing built-in data storage. In order to be considered a mobile device, the device must host an operating system which is not considered a complete operating system such as found on a desktop or laptop and have compatible applications which are available through multiple means. The smartphone has been described by Androulidakis (2016) as one of the most characteristic digital devices of current times, with its pervasion into everyday life and worldwide distribution.

Over time, an abundance of data is collected and stored on a smartphone based on its use, which is strongly correlated to the device's owner or primary user. This stored data has become increasingly sought after and may be involved in court trials to support the solving of crimes with evidence which can be retrieved from the device. Some types of evidence which can be extracted from a mobile phone as identified by Androulidakis (2016), include; device location, which can be based on the serving base location or GPS data; association of

contacts, based on call logs which detail incoming and outgoing calls to and from the device; and communication content such as messages and emails which may have been sent, received or stored on the handset. Digital Evidence has been defined by the National Institute of Justice (2016) as information which is stored or transmitted in binary form which can be relied on in court. Digital evidence has grown from being primarily associated with electronic crime to now being utilised to support the prosecution of all crimes due to the nature of information which individuals store on their personal devices it may be possible to derive critical evidence regarding criminal intent, whereabouts in relation to a crime, or relationships to other suspects or involved parties.

As identified by Rajendran and Gopalan (2016) in their mobile forensics investigation lifecycle process, once a digital device has been acquired, and information regarding the devices specifications and its surroundings upon acquisition have been gathered, the immediate action which needs to be taken is disabling the device from not only the mobile network and the Internet, but also disabling Bluetooth, tethering, and any other kind of external connections. Androulidakis (2016) states that the main principle of digital forensics is the preservation of data. This supports the idea of data preservation and network isolation as a compulsory measure to ensure that data cannot be altered in order to stand up in court. In the case that digital evidence has been altered, which can occur due to data which can change or destroy the contents of a mobile device being received or transferred, a court case can be lost. Incoming traffic from the network such as phone calls, messages or communication with installed apps can result in the contamination of potential evidence. In addition to this incoming traffic, it is possible that destruction mechanisms could be configured on the device which can result in the deletion of data or the locking of the device (Androulidakis, 2016). The following step in the aforementioned lifecycle is to ensure the preservation of the smartphone's data. It is suggested by Rajendran and Gopalan (2016), in order to support data preservation, that electronic devices be stored in a Faraday bag.

Faraday bags are described by Doherty (2016) as similar in appearance to antistatic bags, with the difference being that an antistatic bag will prevent damage to electronic devices from static electrical charges which have built up, whereas a Faraday bag poses the purpose of protecting electronic devices from external connectivity. Faraday bags are based on the concept of the Faraday cage, which is an enclosure which prevents external signals from reaching electronic devices. Doherty (2016) details the effectiveness of a Faraday bag being reliant on the materials of which it is made, the purpose of which is to stop wireless signals from penetrating the bag and reaching the device, which in turn supports the protection of the devices integrity from external influences. While awaiting examination, it is crucial that effective measures are taken which prevent any data which is stored on a device from being altered or remotely destroyed. Gershowitz (2013) identifies three operations which could be utilised in order to protect device integrity without utilising changes to the running applications so as switching off WiFi or entering Airplane Mode, these are: an extraction device that can copy phone contents to a secure offsite location; a Faraday bag; or alternatively a piece of aluminium foil. The benefits of the data extraction device at the time of seizure are evident. However, it is likely that high costs will be associated with such a device and it is unlikely that a forensic investigator or Police Officer will have such a device with them at all times. The benefit of using a Faraday bag is that once inside, a phone is no longer able to communicate with any external sources, and in turn, external sources are unable to reach a phone which is stored in a Faraday bag.

Generally, the primary material which Faraday bags are made from is aluminium foil, which suggests that the utilisation of a simple solution such as wrapping a mobile device in aluminium foil may provide many of the same benefits as a Faraday bag. Although it is possible that aluminium foil may not completely prevent all attempts at communication with the wrapped device, it may act as a suitable intermediary process between the seizure of the device and the secure transportation and storage of a device until further actions can be taken with the potential evidence. In support of the Faraday option Androulidakis (2016) suggests that the best solution for isolating a device from the network and preserving its data is the use of a "Faraday Cage" which isolates electromagnetic radiation. It is stated that Faraday bags are smaller, more convenient option for device isolation. Ayers et al. (2014) define evidence preservation as the process for securely maintaining custody of a digital device without its contents being altered, and is the first step in digital evidence recovery. Incorrect procedures, or improper handling of a device can cause the loss or alteration of digital data. An inability to correctly preserve evidence can forfeit a whole investigation, and potentially lead to the loss of a legal case due to the acquired evidence being disrupted and therefore not standing up in court (Ayers et al., 2014). NIJ (2008) detail recommended items which first responders should have to comprise a "digital evidence collection toolkit" to assist them in performing their investigation in a forensically sound manner. Within this toolkit, alongside notepads, gloves, and a camera for documenting evidence, it is suggested that radio frequency shielding materials be on hand in the case that smartphones or other mobile communication devices are involved in an investigation and need to be securely seized. The specific radio frequency shielding materials which are suggested by NIJ (2008) are Faraday isolation bags, or aluminium foil.

NIJ (2008) emphasise the importance of leaving a mobile device in the power state in which it was found. If the phone is on when the first responders arrive, the phone should be left on, if it is powered off, then it should remain off; the device should then be packaged in a material which will shield it from incoming signals, in preparation for secure transportation. In contrast to this, SWDGE (2011) state that in the event a mobile device cannot be processed immediately, it should be powered off, its battery removed, and not turned back on. Benefits of turning mobile devices off include: the preservation of call logs and last cell tower location information (LOCI); avoiding overwriting of deleted data; stopping remote data destruction signals from reaching the device; and preventing accidental device usage such as messaging, dialling, and accessing and altering files and data. The risks of switching off the mobile device can result in the activation of security and authentication mechanisms such as PIN codes and passwords which further restrict access to the devices content (SWDGE, 2011). However, in support of NIJ's (2008) claims, SWDGE (2011) state that in the event a mobile device must remain powered on, it should then be isolated from the network. Rather than suggesting the use of radio frequency shielding material to avoid the mobile device communicating with cell towers and in turn, altering the phones data, SWDGE (2011) propose that mobile devices can be switched to "Airplane" mode to limit their access to the towers, and where practical, first responders should disable the device's WiFi, Bluetooth, RFID, and infrared communications. Whilst the guidelines provide for a range of procedures which at times are conflicting, it was decided for the purpose of these experiments that a forensic investigator would likely place the phone in a Faraday bag or other shielding material without altering any settings or switching off any technologies.

There is currently a paucity of empirical research on the effectiveness of common Faraday bags for mobile signal shielding. There was even less available in regards to literature which provided findings, results or suggestions in relation to alternative materials which may be utilised in the place of Faraday bags for the purpose of blocking mobile signals from reaching a smartphone. That which could be derived from the literature surrounding alternative materials is the effectiveness of tinfoil regarding the blocking of signals can be attributed to its conductivity. This suggests that any conductive material such as copper, iron, and steel have the potential to create a barrier between a mobile device and radio signals or at least cause interference between the sending and receiving of frequency signals (Barrett, n.d.; Science Buddies, 2011).

RESEARCH DESIGN

The first step in the design of the experiments was to review several previous studies into shielding properties of various materials. None of these studies provided all necessary steps that were suitable for this research so ideas for testing were based on the previous research but with necessary modifications. As mobile devices may have up to three different wireless technologies active, 3G / 4G, WiFi and Bluetooth, it was decided to test all three. It was recognized that different phone manufacturers and models may provide differing results so six common phones from three different manufacturers were chosen for the tests. Five easily available Faraday bags were chosen for testing representing four different manufacturers. To compliment these dedicated Faraday bags and to act as a comparison, tin foil and a tin can were added to test as these were materials that could likely be located by an investigator at a scene when a Faraday bag is not available. The experimental design is described in the following section.

Experimental Design

This research investigates the shielding and preservation capabilities of a range of different Faraday bags, with the main research question to be addressed being: *What is the capability of Faraday bags and alternative materials for blocking mobile network, Wi-Fi, and Bluetooth signals to mobile devices for the purpose of data preservation?* This research provides an opportunity to explore the effectiveness of materials in supporting mobile evidence preservation in order to improve the mobile forensic investigation process. By providing a comparison of a range of available Faraday bags, in addition to two alternative household materials which could be used in the absence of Faraday bags, forensic investigators have the capability to identify materials which may be beneficial to their investigations, and provide support to the preservation of potential evidence.

Table 1: Experimental phones & materials

Mobile Devices		Shielding Materials
Samsung Galaxy S3		*Faraday Defence 3mm*
Samsung Galaxy S5	*Tested with*	*Faraday Defence 7mm*
Samsung Galaxy S7		*EDEC Black Hole Faraday Bag*
Sony Xperia Z5		*Blackout Faraday Shield*
Apple iPhone 5		*ESD Faraday Cage Bag*
Samsung GT-B2710		*Aluminium foil & Tin can*

Research Method

To ensure consistent results, the same method was implemented for each mobile device and Faraday bag combination. The prerequisites for testing included:

- A fully charged battery, due to the high drain nature of the testing
- Device volume set to maximum to support identification of failed shielding where bags restrict vision and interaction
- Installed and configured Viber application for placing and receiving calls through Wi-Fi connection
- Bluetooth turned on, device visibility set as discoverable, and discovery timeout set to "Never"
- Screen timeout set to the longest period available per device
- Mobile Data turned off as to not interfere with the Wi-Fi tests

Prior to shielding the receiving device, two calls will be placed, one over the network, and one over Wi-Fi, using the Viber application, to ensure that all required features were functioning appropriately. In addition, a Bluetooth search will be conducted from the transmitting device to ensure visibility of the phone being tested.

The mobile device is placed in a Faraday bag and a timer pre-set for 30 second intervals over the duration of two minutes. Within every 30 second interval, a network call, and a Viber call is made from the transmission device, to the shielded, reception device. If the communications are received by the shielded device, the test will be considered a failure, and in the event no signals are received then a success is recorded. In the event the majority of results for a specific shielding bag have failed, the shielded device is placed within another bag, so that the device is nested within two bags of the same make and the testing process is then repeated. If after two bags the results are still mostly failures, the testing process is performed again with three nested bags. If after three nested bags the tests still fail, the Faraday bag in question will be considered completely ineffective.

Independent tests are conducted for Bluetooth connectivity to determine the effectiveness of the shielding materials at varying distances, 1 metre, 3 metres, 5 metres and 10 metres. In the event of a failed test, the nesting approach will not be implemented for the Bluetooth testing but instead the distance between the transmitting and shielded receiving device will be increased. If after 10 metres Bluetooth communication can still be established, the material being tested will be considered ineffective.

RESULTS

The 7 shielding materials, 5 Faraday bags, tin foil and a tin box were all tested using similar parameters. Of primary interest is the 3G / 4G shielding capabilities as this is the communication technology that provides the capability to alter the data on the phone from a considerable distance. The results of the experiments are shown in figure 1 where the vertical bar represents a failure to shield the reception.

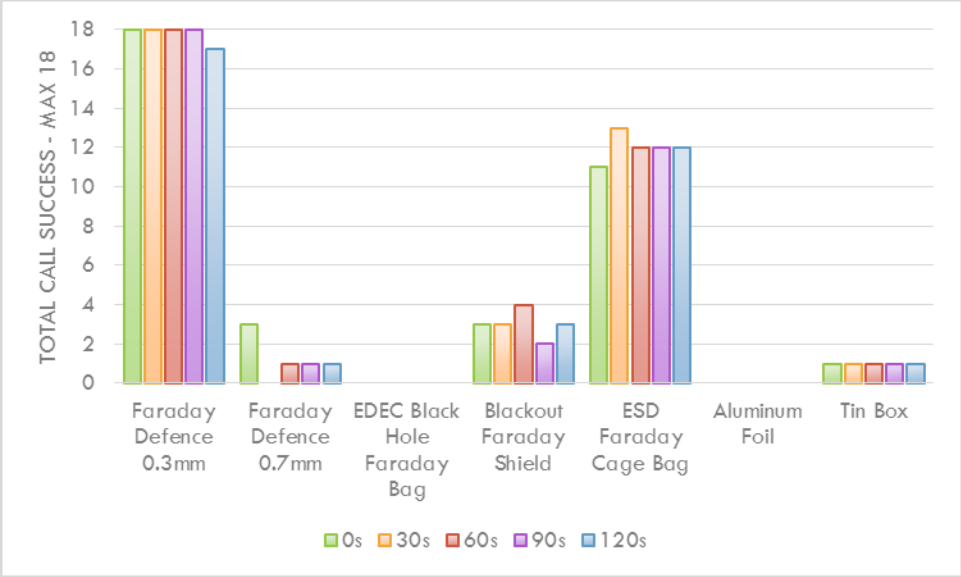


Figure 1. 3G / 4G Call Failures

Of the 7 tested scenarios, only the EDEC Black Hole Faraday Bag and the aluminium foil shielded all 18 calls from penetrating to the device. The time of the phone call had minimal impact on the measured ability of the material so that a very short call is a good indication of whether the material is effective or not. Whilst the Faraday Defense 3mm performed the worst of the 7, 4 other materials let through at least one of the calls. Surprisingly, tin foil performed equally to the best Faraday Bag with no calls getting through to the device. The next communication technology tested is WiFi as this technology has potential communication distances of several hundreds of metres. The results for the WiFi tests are shown in figure 2.

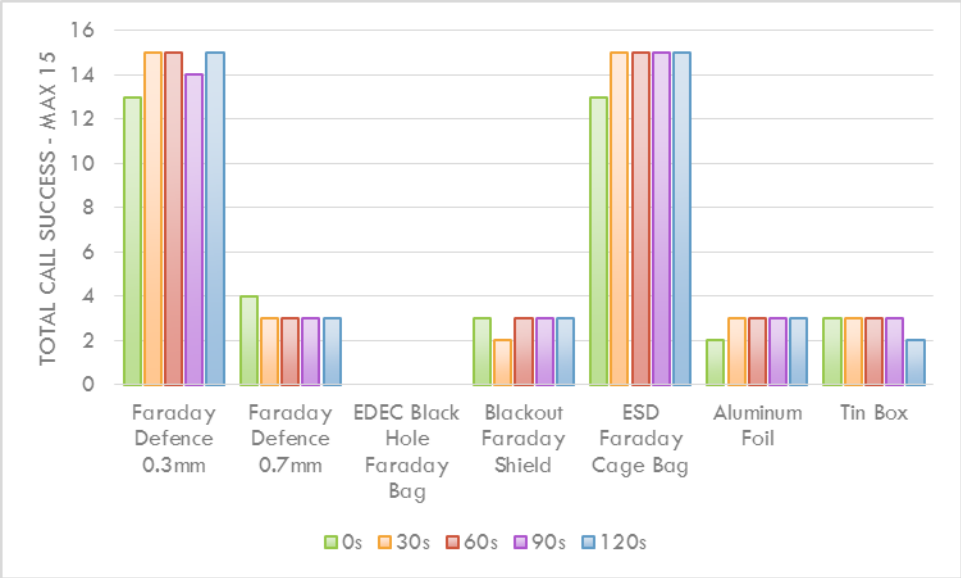


Figure 2. WiFi Call Failures

The 2.4GHz WiFi tests show that the most effective material is the EDEC Black Hole Faraday Bag with none of the communication attempts able to penetrate the bag. The other materials let through at least 2 of the attempts at communication with aluminium foil allowing as many attempts as several of the other scenarios. The final communication technology tested is Bluetooth and the results for this experiment are shown in figure 3.

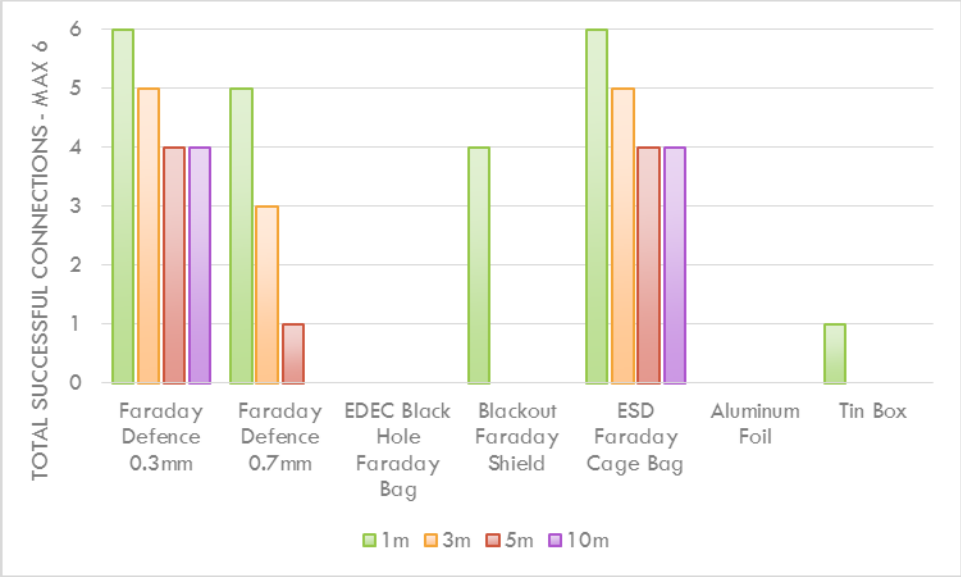


Figure 3. Bluetooth Call Failures

For the Bluetooth tests, the distance between the devices has a noticeable impact on the effectiveness of the materials and this can be expected with the very limited range of Bluetooth. The following tables 2 – 6 show the results for the various phone models with the 5 Faraday bags. Ticks indicate that the signal has penetrated the bag where a cross indicates a successful shielding of the signal.

Table 2: 3G / 4G and WiFi results for Faraday Defence 0.3mm

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	✔	✔	✔	✔	✔		x	✔	✔	✔	✔
Samsung Galaxy S5	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Samsung Galaxy S7	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Sony Xperia Z5	✔	✔	✔	✔	✔		x	✔	✔	✔	✔
Apple iPhone 5	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Basic Phone	✔	✔	✔	✔	✔		**n/a**	**n/a**	**n/a**	**n/a**	**n/a**

Table 3: 3G / 4G and WiFi results for Faraday Defence 0.7mm

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	x	x	x	x	x		✔	✔	✔	✔	✔
Samsung Galaxy S5	✔	x	x	x	x		✔	✔	✔	✔	✔
Samsung Galaxy S7	x	x	x	x	x		✔	✔	✔	✔	✔
Sony Xperia Z5	x	x	x	x	x		x	x	x	x	x
Apple iPhone 5	✔	x	x	✔	✔		✔	x	x	x	x
Basic Phone	x	x	✔	x	x		**n/a**	**n/a**	**n/a**	**n/a**	**n/a**

Table 4: 3G / 4G and WiFi results for EDEC Faraday Bag

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	x	x	x	x	x		x	x	x	x	x
Samsung Galaxy S5	x	x	x	x	x		x	x	x	x	x
Samsung Galaxy S7	x	x	x	x	x		x	x	x	x	x
Sony Xperia Z5	x	x	x	x	x		x	x	x	x	x
Apple iPhone 5	x	x	x	x	x		x	x	x	x	x
Basic Phone	x	x	x	x	x		**n/a**	**n/a**	**n/a**	**n/a**	**n/a**

Table 5: 3G / 4G and WiFi results for Faraday Cage ESD Bag

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Samsung Galaxy S5	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Samsung Galaxy S7	x	x	x	x	x		✔	✔	✔	✔	✔
Sony Xperia Z5	x	x	x	x	x		x	✔	✔	✔	✔
Apple iPhone 5	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Basic Phone	✔	✔	✔	✔	✔		n/a	n/a	n/a	n/a	n/a

Table 6: 3G / 4G and WiFi results for Blackout Faraday Shield

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔
Samsung Galaxy S5	x	x	x	x	x		x	x	x	x	x
Samsung Galaxy S7	x	x	x	x	x		✔	✔	✔	✔	✔
Sony Xperia Z5	x	x	x	x	x		✔	x	✔	✔	✔
Apple iPhone 5	✔	✔	x	x	✔		x	x	x	x	x
Basic Phone	✔	✔	✔	✔	✔		n/a	n/a	n/a	n/a	n/a

Of particular interest in these results is that whilst only the EDEC Faraday bag is successful in all experiments, other bags have varied success depending on the mobile phone make and model, which different models from the same manufacturer providing different results. Tables 7 and 8 show the results for the aluminium foil and tin can as a shielding material.

Table 7: 3G / 4G and WiFi results for Aluminium Foil

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	x	x	x	x	x		x	x	x	x	x
Samsung Galaxy S5	x	x	x	x	x		x	✔	✔	✔	✔
Samsung Galaxy S7	x	x	x	x	x		x	x	x	x	x
Sony Xperia Z5	x	x	x	x	x		✔	✔	✔	✔	✔
Apple iPhone 5	x	x	x	x	x		✔	✔	✔	✔	✔
Basic Phone	x	x	x	x	x		n/a	n/a	n/a	n/a	n/a

Table 8: 3G / 4G and WiFi results for Tin Box

	3G/4G						Wi-Fi				
	0s	30s	60s	90s	120s		0s	30s	60s	90s	120s
Samsung Galaxy S3	x	x	x	x	x		x	x	x	x	x
Samsung Galaxy S5	x	x	x	x	x		✔	✔	✔	✔	✔
Samsung Galaxy S7	x	x	x	x	x		✔	✔	✔	✔	✔
Sony Xperia Z5	x	x	x	x	x		x	x	x	x	x
Apple iPhone 5	x	x	x	x	x		x	x	x	x	x
Basic Phone	✔	✔	✔	✔	✔		n/a	n/a	n/a	n/a	n/a

Table 7 indicates that aluminium foil can be utilised for 3G / 4G signal shielding and is successful at shielding the Sony Xperia Z5 and Iphone 5 from WiFi signals. In contrast, the tin box will prevent signal penetration of 3G / 4G for the basic phone which is not equipped with WiFi and will shield against WiFi for the Samsung Galaxy S5 and S7 only. These results clearly show that no one rule for shielding material can be applied in all circumstances with all phones or all technologies.

CONCLUSION

The use of Faraday bags, and in exceptional circumstances other materials to prevent unwanted radio communications with seized wireless devices is an accepted and vital addition to the forensic investigators toolkit. Forensic evidence is only of evidentiary value if it is maintained in the state that it was seized in without modification or alteration. The ready availability of Faraday bags at little cost would seem to allow investigators

to cheaply equip themselves with Faraday bags in readiness for a mobile phone seizure. However, the results of this research show that only one of the 5 tested bags consistently provided reliable protection from radio waves penetrating the bag. This same bag provided radio wave blocking for all 3 tested communication technologies and surprisingly aluminium foil was as reliable in all but blocking of WiFi signals. The other bags, even when nested inside each other to provide greater protection failed to completely block the radio signals with selections of phones but that only testing of the phones and materials provided a guide as to their effectiveness. It is not sufficient to estimate the effectiveness based on how they perform with some phones and extrapoloate the results to incorporate untested phones. The forensic investigator should be aware that not all bags are of equal reliability and the choice of which bag to utilise from which manufacturer should be made after reviewing rigorous testing procedures such as these to ascertain their effectiveness. Additionally, should a reliable bag not be available, aluminium foil will serve as an emergency measure provided the WiFi on the device is disabled by the investigator.

REFERENCES

Androulidakis, I. I. (2016). *Mobile phone security and forensics : A practical approach* (2 ed.). Retrieved from http://AUT.eblib.com.au/patron/FullRecord.aspx?p=4455173

Barrett, J. T. (n.d.). *Why does aluminium foil block cell phone signals?* Retrieved from http://techin.oureverydaylife.com/aluminium-foil-block-cell-phone-signals-2475.html

Bennett, D. (2012). The challenges facing computer forensics investigators in obtaining information from mobile devices for use in criminal investigations. *Information Security Journal: A Global Perspective, 21*(3), 159-168. doi:10.1080/19393555.2011.654317

Doherty, E. P. (2016). *Digital forensics for handheld devices* (1 ed.). Retrieved from http://AUT.eblib.com.au/patron/FullRecord.aspx?p=981555

FireEye. (2015). *Out of pocket: A comprehensive mobile threat assessment of 7 million iOS and Android apps*. Retrieved from https://www2.fireeye.com/MobileThreatAssessment.html

Gershowitz, A. M. (2013). Seizing a cell phone incident to arrest: Data extraction devices, Faraday bags, or aluminium foil as a solution to the warrantless cell phone search problem [article]. *The William and Mary Bill of Rights Journal, 22*(2), 601-612.

Khan, J., Abbas, H., & Al-Muhtadi, J. (2015). Survey on Mobile User's Data Privacy Threats and Defense Mechanisms [Article]. *Procedia Computer Science, 56*, 376-383. doi:10.1016/j.procs.2015.07.223

Ophoff, J., & Robinson, M. (2014). *Exploring end-user smartphone security awareness within a South African context.* presented at the meeting of the 2014 Information Security for South Africa, doi:10.1109/ISSA.2014.6950500

Rajendran, S., & Gopalan, N. P. (2016). Mobile Forensic Investigation (MFI) life cycle process for digital data discovery (DDD). *Proceedings of the International Conference on Soft Computing Systems (ICSCS) 2015, 2*, 393-403. doi:10.1007/978-81-322-2674-1_37

Science Buddies. (2011). *Block radio waves*. Retrieved from http://www.scientificamerican.com/article/bring-science-home-block-radio-waves/

Soukup, P. A. (2015). Smartphones. *Communication Research Trends, 34*(4), 3-39.

AN EXPLORATION OF ARTEFACTS OF REMOTE DESKTOP APPLICATIONS ON WINDOWS

Paresh Lalji Kerai, Vimal Murji Vekariya
Security Research Institute & School of Science
Edith Cowan University, Perth, Australia
p.kerai@ecu.edu.au, vvekariy@our.ecu.edu.au

Abstract

Remote Desktop Applications (RDA) such as Virtual Network Computing (VNC), Cisco WebEx, GoToMeeting and LogMeIn have been adapted and utilised recently. This is because they facilitate tier-one support to configure computers, networks and solve application-related issues from a remote location. The direct benefit from the use of these applications, is the time (and therefore cost) saving for organisations. Unfortunately, "remoting" technology can also be used by criminals to perform illegal activities, hence remote applications are of key interest to law agencies and forensic investigators. The research outlined in this paper aims to identify any artefacts left behind by common remote applications and technologies used by many firms. These artefacts can be vital to government law enforcement agencies and forensic investigators, as they could be used as evidence in cyber-crime investigations. This research will focus on RealVNC, TightVNC, Cisco WebEx, GoToMeeting and LogMeIn applications. The findings from the research shows some artefacts left behind by the applications, which can be used by forensics investigators or law enforcement for possible evidence.

Keywords
Virtual Network Computing, Computer Forensics, Encryption, Computer Security, Remote Desktop

INTRODUCTION

RDA allow organisations and individuals to access a computer or network remotely, where physical access to the network is not available or where access to the device is impractical. Such applications facilitate access to computers as though a user is sitting right in front of his/her computer. There are various types of RDA currently used by many organisations and individuals such as VNC, LogMeIn, Citrix GoToMyPc, LogMeIn and TeamViewer. The increasing adoption and use RDA has also increased the potential for cyber threats to be realised against organisations

However, using the application can be slow over the internet depending on the Internet connection speeds. Some remote applications such as VNC have had security issues in past due to using weaker encryption implementations (Kerai, 2010). It is plausible that similar issues exist in other remote applications, or still exist in VNC.

In this paper, we detect and retrieve any artefacts left behind by RDA on a Windows computer. Client software for five common remote applications (RealVNC, TightVNC, LogMeIn, Teamviewer and Citrix GoToMyPc) will be installed individually on a Windows 10 computer, and test remote connections will be initiated that we expect will create configuration changes on the computer. The information will be forensically analysed to find any artefacts left by the application on the computer. The paper outlines the location of the artefacts and the types of information left by the application.

RealVNC, TightVNC, LogMeIn, Teamviewer and Citrix GoToMyPc remote applications was used for to conduct the research. The findings will show whether the R leave any artefacts, and this can help law enforcements for forensic investigations.

BACKGROUND

Computer forensics is computer science used to aid legal process in investigations (Brown, 2006). It is the process of obtaining, identifying, extracting, analysing and documenting digital evidence for use as evidence in legal case (Brown, 2006; Kerai, 2010). However, forensics analysis need to be followed in a defined procedure which is accept by the law court. A guideline defined in HB171 (HB171, 2003) is currently used as a guide to carry out forensics investigations and manage electronic evidence.

It is important to follow a standard procedure when performing acquisition of electronic evidence when performing a forensic investigation (Hannay, 2008). It is important to acquire the evidence without modifying or damaging it and validating the evidence is the same as original. This is done by chain of custody and documentation.

Forensic analysis for remote protocols has been an area of interest for the law enforcements and other government and state agencies. This is because these applications can be used to also perform cyber-crimes and illegal communications. A previous research done by (Kerai, 2010), showed that VNC and Microsoft Remote Desktop Protocol left behind artefacts, connection and application logs on a computer system locally. The research showed that the VNC application and Remote Desktop Protocol leaves some artefacts on the Windows Registry File System and other also connection and other application logs on the computer system locally.

REMOTE DESKTOP APPLICATIONS

Demand for RDA has increased with time, where a user can remotely connect, manage and configure another computer or networks. Connectivity between a user's computer to another remote computer can be achieved with readily available hardware and software to connect you virtually to any remote network. The constraint is cost and bandwidth issue that is associated with some technologies (Hoogenboom & Steemers, 2000).

There are various remote desktop applications and technologies available in recent years that can be used to remote in and fix a computer related issues. Online tools like LogMeIn and Citrix GoToMyPc are easy to use, reliable and secure (TeamViewer, 2016). However, standalone applications such as RealVNC, TightVNC, UltraVNC, Teamviewer and others are applications will require the user to configure the application on the computer for a remote connection, which may require some technical skills.

VNC (Virtual Network Computing) - VNC was first developed by the Olivetti and Oracle Laboratory as their telephony system. Later the technology was acquired by AT&T labs and the owners of the technology formed RealVNC to continue working with the remote technology.

The VNC application uses a Remote Frame Buffer protocol to remote access a graphical user interface of the connected device. The application uses TCP 5900 and TCP 5800 for web based remote access. Hence it enables users and organisation to access network resources remotely over the internet. The application has two independent versions, the client and the server, both versions run on most operating system platforms. This makes it very popular as anyone using Windows operating system can remote in to a Linux-based computer or Mac OSX based computer and vice versa. However, a server instance needs to be installed on a computer and a client is then used to access the server. Previous research has shown and discussed the results on how the VNC connection works and the security features and weakness it has (Kerai, 2010).

Other remote support applications such as LogMeIn, GoToMyPc and Teamviewer use TCP port 443 and have different application architecture and implementation than VNC application, due to them being web client based rather than standalone. Hence the artefacts might be different compared to VNC application.

Next session outlines the lab setup and materials used to analyse the remote applications. The research outlines artefacts left by the applications and individual or remote connection information. These artefacts and information can be used by forensics investigators and law agencies in court of law as evidence, in a cyber-crime related criminal case.

EXPERIMENTAL SETUP

Various software and tools were used during the data collection, acquisition and analysis of the raw dd image of the Windows operating systems (see table 1). Different virtual machines and images were used to analyse the application independently.

Table 1: Tools used in Experiments

Applications	
VMware Workstation v12 professional	Windows 10 operating system was installed and created separate instances of each application for analysis.
Remote Desktop Applications	• Real VNC version 5.3.2 • TightVNC version 2.7.10 • Citrix GoToMyPC • Teamviewer • LogMeIn
Password Recovery Tools	• Abel and Cain v 4.9.35 – This is a password recovery and sniffing tool used to capture network traffic and sniff passwords. • VNCPass • VNCSniff
Windows File System Tools	• Microsoft Sysinternal Suite

ANALYSIS

All RDA were independently analysed on a virtual machine running the Windows 10 operating system. An ADSL router by default does not have remote connections to computers externally, this is because the router does not have remote protocol ports open on the router firewall. For VNC connection to work the user needs to open TCP port 5900 and TCP port 5800 for the Java client. However other mentioned remote applications run on TCP port 443, therefore they do not need special ports opened on the router. Once a VNC application is installed on a computer it automatically opens the ports on the local Windows firewall.
VNC applications use Data Encryption Standard (DES)which has proven to be insecure due to its small key size (56-bit). Due the weak encryption algorithm, an attacker can sniff the traffic after cracking the encrypted password.

Windows Registry Analysis

The Windows registry is a central hierarchical database used by the Windows Operating System, to store all the information that is necessary to configure and manage applications and hardware devices installed on the system (MicrosoftSupport, 2016). The registry hive is a group of root keys and sub keys that contain application data (Alghafli, Jones, & Martin, 2010). There are five logical hives in the Windows Registry, the application registry values are most stored under the HKEY_CURRENT_USER, HKEY_LOCAL_MACHINE and HKEY_USERS registry hives.

RealVNC

RealVNC configuration is stored in the Windows registry, including encrypted login passwords, settins and outbound connections made to the RealVNC servers. The default Windows registry viewer was used to analyse the data and artefacts of the application. The connection server password used by the RealVNC server is encrypted with DES encryption, since DES encryption is weak, the password can be retrieved easily. Tools such as Cain and Abel, VNC Crack and online password cracking web services can crack DES encryption with no cost, due to readily available tool to decrypt the encryption. Table 2 shows the artefacts left by RealVNC application on Windows 10 machine.

Table 2: RealVNC Windows Registry Analysis

Under HKEY_CURRENT_USER\SOFTWARE\RealVNC\vncviewer\MRU, the application stores all the history of the external IP addresses connections made to an external VNC server.
Under HKEY_LOCAL_MACHINE\SOFTWARE\RealVNC\vncserver, the application stores encrypted password of the RealVNC server. The password is encrypted with DES encryption standard. Since DES is a weak encryption standard, decrypting the password is possible.
Under HKEY_LOCAL_MACHINE\SOFTWARE\RealVNC\vncserver, the application also stores configuration settings such as, if the application is personal or enterprise edition, RSA private key and the authentication scheme used.
Under the HKEY_USERS\Software\RealVNC\vncserver, the application stores the vncserver license key.

Table 3: RealVNC Logs on Windows File System

• The application leaves log information in the Windows Event Viewer under the Application Logs. The logs show client IP address that made the connection to the RealVNC server, and the connection logs for the computer made a connection to another VNC server.
• The RealVNC application stores the connections and application log settings on the hard drive. The log file includes all the connection details, including the client IP address and any files transferred during the connection. The file is located under C:\Program Files\RealVNC\VNC Server\Logs or C:\Program Files (x86) \RealVNC\VNC Server\Logs. However, by default the application logging is not enabled, therefore a user will need to enable logging for the logs to be generated.
• By default, Windows Firewall logs are not enabled, therefore no log information is kept if a VNC connection is made inbound or outbound. However, if the Windows firewall logging is enabled, then the VNC connection is logged on the log file. The firewall log file is located under C:/Windows/System32/LogFiles/Firewall
• The application stores the chat conversation logs under the folder C:\Users\admin\AppData\Local\RealVNC.

TightVNC

TightVNC is a free application, and works the same way as RealVNC. The application also stores the application settings and logging information on both the Windows Registry structure and locally on the computer hard drive. The connection server password used by the TIghtVNC server is encrypted with DES encryption, in a similar way to how RealVNC server encrypts its password. As stated early the encryption is weak and can be defeated easily to retrieve the server password. The table below shows the artefacts left by RealVNC application on Windows 10 machine.

Table 4: TightVNC Windows Registry Analysis

• Under HKEY_CURRENT_USER\SOFTWARE\TightVNC\Server, the application stores the configuration settings for the TightVNC server, this includes also the encrypted password for the server. This password is used to connect to the server. • Just like RealVNC the connection password is encrypted with DES and therefore password can be attacked with password cracking tools.
• Under HKEY_CURRENT_USER\SOFTWARE\TightVNC\Viewer, the application stores all the connection history of the local computer making VNC connections outbound. Therefore, the registry key stores all the IP addresses the TightVNC server has made connection to another external VNC server.
• Under HKEY_LOCAL_MACHINE\SOFTWARE\TightVNC\Server, the application also stores the configuration settings for the TightVNC server, this includes also the encrypted password for the server. This password is used to connect to the server.

Table 5: TightVNC Logs on Windows File System

• The application leaves log information in the Windows Event Viewer under the Application Logs. The logs show client IP address that made the connection to the TightVNC server, and the connection logs for the computer made a connection to another VNC server.
• The application stores the connections and application log settings on the hard drive. The log file includes all the connection details, including the client IP address and any files transferred during the connection. The file is located under C:\ProgramData\TightVNC\ Server\Logs. However, by default the application logging is not enabled, therefore a user will need to enable logging for the logs to be generated. • Also, the application does not log any outgoing VNC connections, therefore if a user makes a VNC connection to an external VNC server, no logging information is kept by the application.
• By default, Windows Firewall logs are not enabled, therefore no log information is kept if a VNC connection is made inbound or outbound. However, if the Windows firewall logging is enabled, then the VNC connection is logged on the log file. The firewall log file is located under C:/Windows/System32/LogFiles/Firewall

Citrix GoToMyPc

Citrix GoToMyPc is an application that can be used to access remote computers over internet connection. Unlike VNC applications, GoToMyPc application does not require client and server, allowing invitations to other users to remotely access their computer. The connection is done over HTTPS 443 TCP port; therefore, it does not require to have an independent port open on the firewall to work.

The user initially will need to register on the Citrix GoToMyPc website and add remote computers to the registered account. However, the remote computers need to have a GoToMyPc installation to receive the remote connections. User will need to log in to the account and then initiate a remote connection by selecting the computer defined on the user account. Someone can also send out an invite to connect a computer that is not registered to the account and initiate a connection if accepted by the guest user. The application uses end to end encryption during the remote connection using the AES 128bit encryption standards (GoToMyPc, 2016). The table below shows the artefacts left by RealVNC application on a Windows 10 machine.

Table 6: GoToMyPc Windows Registry Analysis

• Under HKEY_LOCAL_MACHINE\WOW6432Node\Citrix\GoToMyPc, the application stores all the configuration settings. The settings include encrypted access code with Advanced Encryption Standard (AES) 128-bit encryption standard, email address of the person the account is registered under and other various settings.
• The application also keeps are the records of guest invites send to connect to the GoToMyPc workstation. These settings can be located under HKEY_LOCAL_MACHINE\WOW6432Node\Citrix\GoToMyPc\GuestInvite.
• Under the registry value HKEY_CURRENT_USER\SOFTWARE\Citrix\GoToMyPc\FileTransfer\history and HKEY_USERS\S-1-5-21-97110503-761733263-3747825532-1001\SOFTWARE\Citrix\GoToMyPc\FileTransfer\history, the application stores the hostname of the computer made the remote connection, including the location of any files transferred during the remote session process.

Citrix GoToMyPc application did not log any access or connection logs on the local computer. The application instead stores the connection logs on the user's GoToMyPc online account. Hence no logs were available on the local computer of any type of connections made or received. Also, no connection or applications logs were stored on the Windows Event Viewer logs.

Teamviewer

Teamviewer is an organisation founded in 2005 in Germany that provides web based collaboration solutions and remote desktop capable support to the users. The application can be used by users to remotely control another computer that has the application installed on it. The application works over TCP port 443 and uses RSA public/private key exchange and AES 256-bit session encryption to secure the connection traffic (TeamViewer, 2016). The table below outlines artefacts left behind by the application on the Windows Registry File Structure.

Table 7: Teamviewer Windows Registry Analysis

• Under HKEY_CURRENT_USER\SOFTWARE\Teamviewer; registry key value, it stores the application settings including the account name of the Teamviewer account. This is the email address used to login on to the Teamviewer account. The application also stores the username of the local computer.
• Under HKEY_LOCAL_MACHINE\WOW6432Node\TeamViewer and HKEY_USERS\S-1-5-21-97110503-761733263-3747825532-1001\SOFTWARE\Teamviewer; registry values, the application stores the all the encryption keys and certificates that is generated, when a remote connection is initiated and other Teamviewer settings.

Table 8: Teamviewer Logs on Windows File System

• The application stores the incoming connection logs under the C:\Program Files (x86) \TeamViewer\Connections_incoming, the file includes all the incoming connections to the computer. However, Teamviewer doesnot use an IP address or email address for initiating a remote connection, instead it uses a unique random generate number, which is used as a ID to initiate a connection.
• Under the same folder C:\Program Files (x86) \TeamViewer\, Teamviewer stores a log file named "TeamViewer11_Logile", this file contains all the remote connection information, including the client ID, the public IP address of the client that made the connection, log entries for any files transferred during the remote session. This is crucial for the forensic investigation as it can be used to track down an individual with a Public IP address.

The application stored no connection or applications logs were stored on the Windows Event Viewer logs.

LOGMEIN

LogMeIn provides remote connectivity solutions to organisation for collaboration and IT management. The organisation was found in 2003 and provides servers to user to access and connect to remote computers. The LogMeIn client initially establishes a connection to LogMeIn servers and authenticates itself with a TLS1.X certificate. Once the identity is verified the user can then connect to and exchange data to another host to computer associated to the LogMeIn account. The application uses TLS based certificate authentication and uses AES or Triple Data Encryption Standard (3DES) encryption 128-bit or 256-bit encryption standards (LogMeIn, 2016). The table below outlines artefacts left behind by the application on the Windows Registry File Structure.

Table 9: LogMeIn Windows Registry Analysis

• Under the registry key, HKEY_CURRENT_USER\SOFTWARE\LogMeIn\Toolkit\DesktopSharing, the application stores the last guest recipient email address used to invite a user to remote into the computer.
• Under the registry key, HKEY_CURRENT_USER\SOFTWARE\LogMeIn\Toolkit\Filesharing, the application last guest recipient email address used to share a file transfer from the computer.
• Under the registry key, HKEY_LOCAL_MACHINE\SOFTWARE\LogMeIn, the application stores all the configuration settings and user preferences. The registry value also stores user guest invite settings.
• Under the registry key, HKEY_LOCAL_MACHINE\SOFTWARE\LogMeIn\V5\WebSvc\Shared, the application stores details about any files shared to the remote user.

Table 10: LogMeIn Logs on Windows File System

• LogMeIn application logs all the connection logs both incoming and outgoing to a log files located under the C:\ProgramData\LogMeIn folder. There are two logs files that contain connection settings including the public IP address of the remote user and basic remote connection settings used for the connection. • The log files also contain information about file transfers and send over to remote computer, such as name of the file, location of the file and file type. • The logs also contain details any incoming remote connections initiated to the computer by an external remote user. This includes the public IP address of the computer initiating the remote connection.
• The application incoming and outgoing remote connection is stored and captured on the Windows event viewer, under the Application logs. The log entry stores public IP address of the guest user connecting to the computer. • Once the remote user ends the connection or logged out, another log entry is created capturing the event with the IP address and the time stamp of the connection.

As identified above on the results, the RDA do leave behind useful artefacts such as IP addresses and email addresses of individuals who have remoted in a computer. This information can be used by law agencies and forensic investigators to prosecute or assist them to develop their investigations further. The application incoming and outgoing remote connection is stored and captured on the Windows event viewer, under the Application logs. The log entry stores public IP address of the guest user connecting to the computer. The logs also show information such as file name, type and location of files been transferred to the remote computer. This can be used to find exactly what the kind of files were transferred and the location of the file on the local and remote computer. All the applications store some sort of encrypted credentials and certificates on Windows registry hive. However, VNC applications store server connection password on DES encryption standard, a less secure and can be cracked easily by password cracking tools.

Comparing the results of previous research done by (Kerai, 2010), there is not a much difference on how RDA leave behind artefacts. Especially RealVNC and TightVNC applications do leave behind reasonable number of artefacts including the connection encrypted password used by the applications

CONCLUSION

Remote desktop applications provide end users and organisations the ability to remote in to networks and computers to manage and troubleshooting network related issues. Due to its ease and graphical interface, the adoption of remote access application is rapidly increasing. There are benefits of RDA to organisations and individuals, however the technology can also be exploited by criminals, cyber groups and terrorist groups to perform illegal activities on remote computers and networks.

As shown above, several RDA leave behind local artefacts which can be of importance to forensic investigation. This is a large impact to government law enforcements agencies and forensics investigators to recover any potential artefacts left behind by the applications on the computer for further investigations. The artefacts can be used to investigate crimes committed by individuals who have transferred explicit, illegal and terrorism relates materials and documents over the remote connection.

The research conducted has verified that the remote applications do leave artefacts behind and can be assessed and reviewed in forensically sound manner. Further research will explore and compare other remote applications, along with evaluating the applications use on different operating systems. Providing further important artefacts that can be of a forensic interest.

REFERENCES

Alghafli, K. A., Jones, A., & Martin, T. A. (2010). Forensic Analysis of the Windows 7 Registry*Australian Digital Forensics Conference, held in Perth, Western Australia. Australia,*

Brown, C. L. T. (2006). *Computer Evidence: Collection and Preservation*. Massachusetts: Charles River Media, Inc.

Hannay, P. (2008). *Forensic Acquisition and Analysis of the TomTom One Satellite Navigation Unit*. Proc. xxth Australian Digital Forensics Conference, Perth, Australia.

HB171, S. A. I. (2003). Guidelines for the Management of IT evidence. Sydney: Standards Australia International Ltd.

GoToMyPc, C. (2016). GoToMyPc Technology Security White Paper. Retrieved from https://www.gotomypc.com/downloads/pdf/m/GoToMyPC_Security_White_Paper.pdf

Hoogenboom, M., & Steemers, P. (2000). Security for Remote Access And Mobile Applications. *Computer & Security*, *19*(2), 149-163. Retrieved from http://www.sciencedirect.com.ezproxy.ecu.edu.au/science/article/pii/S0167404800878256

Kerai, P. (2010). *Remote Access Forensics for VNC and RDP on Windows Platform.* Bachelor of Computer Science Honours Edith Cowan University, Perth, Australia.

Kerai, P. (2010). *Remote Access Forensics for VNC and RDP on Windows Platform*. Proceedings of the 8th Australian Digital Forensics Conference, Perth, Australia.

LogMeIn. (2016). LogMeIn Security An In-Depth Look. Retrieved from https://secure.logmein.com/welcome/documentation/EN/pdf/common/LogMeIn_SecurityWhitepaper.pdf

MicrosoftSupport. (2016). Windows Registry Information for Advanced Users. Retrieved

TeamViewer. (2016). TeamViewer Security Information. Retrieved from https://downloadap1.teamviewer.com/docs/en/TeamViewer-Security-Statement-en.pdf

ESTABLISHING EFFECTIVE AND ECONOMICAL TRAFFIC SURVEILLANCE IN TONGA

Brian Cusack, George Maeakafa
AUT University, Christ's University in Pacific
brian.cusack@aut.ac.nz, gmaeakafa@yahoo.com

Abstract

The Pacific Islands are seriously challenged by the growth in wealth and the expansion of international material possessions. On the roads traffic has grown dramatically and the types of vehicles now using Island roads has greatly changed. With the importation of cheap second hand vehicles designed for freeway speeds serious safety issues have grown proportionally with the increasing numbers. In this research we consider the prohibitive costs of traditional traffic controls to economy and propose a light weight highly mobile aerial surveillance system that integrates with ground policing capability. Our research question was: How can road safety and security be enhanced with economical technologies? In addition to collecting and processing live data we have also designed a forensically ready system, and an information system to process the large amounts of data generated by the addition of these technologies into the traffic surveillance processes.

Keywords
Surveillance, Security, Road Safety, Economical Technology, Innovation

INTRODUCTION

Aerial surveillance has a long history of use in the military for monitoring activities on the ground (CambraBascca et al., 2013). Commercially, aerial surveillance is used for monitoring resources such as forests, crops, coast lines, aerial news gathering, and search and rescue. The surveillance activities have relied initially on the human eye and binoculars but now have moved towards sophisticated cameras and digital technologies. These images can be reviewed in real time or can be examined at a later date for forensic analysis and evidence. With regards to traffic surveillance in particular, the increase in the number of vehicles on road has led transport management agencies to seek the use of advanced technologies that can provide better information more cheaply in order to service informed decisions and road safety. This requires collection of precise and accurate information about the state of traffic, conformance, behaviours and road conditions. It is also required to get timely information in case of emergencies (accidents, and so on). In the case of accidents, the time of response is critical in victim survivability and with increased traffic volumes even emergency vehicles struggle to reach incident scenes in a timely fashion. Traditional surveillance methods do not prove to be time effective or cost effective. They can provide useful information about traffic flows but cannot provide useful information regarding traffic flows over a space of time including the vehicle trajectories, routing information and real-time movements (Cheng et al., 2012).

Traffic information is the foundation of traffic management, traffic control, and transportation planning. However, many rural and some urban road segments are not installed with any fixed traffic detectors due to the high cost and historically low volumes. In addition in many of the built-up city areas with high surveillance the motorists have learnt behaviour regarding fixed cameras and mobile vehicles, and have change their behaviour to only comply when a surveillance instrument is present. This is exactly why my research is aiming to see if the using of unmanned aerial vehicles (UAVs or drones) will be effective in the Kingdom of Tonga for improving the flow of information for the control of vehicle traffic. Additionally, mobile traffic sensors are not broadly used, which leads to low traffic information acquisition. As a new tool, UAV or civil drone has its unique advantages of flexibility and mobility, wide view scope, and low cost compared with traditional fixed systems. Hence, in this work, UAVs are introduced to monitor road segments and to identify the problems and challenges that would face a full implementation. At present UAVs have many limitations that require management. Some of the challenges are battery life, attenuation of signals, camera cost, large datasets, lack of training, and other issues that will be documented in this research.

This paper is structured to review the background context, look at visual data processing, consider the big data problem, and to propose an integrated visual surveillance system for traffic management. The research contribution is the problem evaluation and design solution proposed.

PROBLEM BACKGROUND

Tonga has very small islands with over one hundred thousand people in population. Consequently small lightweight UAVs even with their current capabilities of endurance are effective tools for monitoring behaviours on the road (Dijk et al., 2013). Tonga does not have any traffic lights even in the most populated areas and few fixed surveillance cameras. There appears to be a need for some kind of electronic device that will assist the Tonga Police in their traffic duties, that is economically effective and efficient. The objectives for this research are to match the suitability of current off-the-shelf drones under $5000 to the task of traffic surveillance (Cusack & Khaleghparast, 2005); and, then to propose the design for information management system. There are also other practical questions to be considered. For example, Are the weather conditions in Tonga suitable? What about from January to April in the hurricane season? Also there is the consideration of other applications of these lightweight surveillance devices. For example, are they useful for other needs such as searching, Agriculture, Fisheries, Fire department or any other commercial need. These are very important areas in Tonga also. The question is, are we able to adopt some kind of framework that will help our traffic conformance as well as the other needs using the civil drones? Because the most important issue here is the safety of the people and if the civil drone can economically add value they should be more widely used.

Traffic conformance audit on roads is maintained by a mixture of surveillance devices and techniques (Coifman et al., 2004; Goradia & Xi, 2012). A compliant vehicle is one which delivers the transportation service within a profile of measurable attributes that conform to the usage contract. These attributes include behaviour, speed, mechanical safety, registration, licensing and so on. The usage contract has a wide variety of requirements that cover different vehicles, human factors and situations. Traffic intelligence can be gathered from a wide range of sensor networks. The most important aspect of using this intelligence is reference. References are found when for example the speed of the vehicle is compared with the conformance contract. A bus or a truck for example has a different conformance contract than a motorcar. Similarly matters of weight, size, registration, and so on all fall within the spectrum of traffic surveillance and intelligence gathering. Each instance on the road has a different conformance contract and some contracts overlap. The overall result is that any data which is collected has to be assessed against the reference data or template. Consequently the information processing around sensor networks has to be tuned to automate the lower-level decision-making processes and to filter out compliant information (Kastrinaki et al., 2003; Ahmed et al., 2012). Persons deciding how to act require only selective information that is about variations to conformance.

When the information processing processes are automated information may be fed directly to law enforcement or an investigating officer. For example sensors on roads may indicate that a vehicle is overweight or over height or travelling with excessive speed or dangerously (Bakhtari et al., 2006; Amran et al., 2014). For this to happen the primary data has to be processed in such a way that the information provided for decision-makers is summative in nature and represents previous determinations against benchmarks. To achieve this level of sophistication highly intelligent information systems are required to support a sensor network. This would mean that when adding UAVs to a surveillance network that the extra intelligence data can be processed into useful information for the decision-makers effectively and efficiently. The software and the computational algorithms required have to be built into the information system to support the higher-level decision-making capability. In practical terms this means that a match between an image of a vehicle and a database is required in order to find information such as the current warrant fitness, the current owner and other useful information that can lead to those responsible for the conformance contract.

VISUAL DATA COLLECTION

A quad-copter drone was used to take video footages of vehicles travelling over city streets. These video footages were then used as the input of the data processing life cycle that was used to automatically process number plates and speed data. The data could be processed either on the drone motherboard, stored for later analysis or in our case we beamed directly back in real time to a laptop processing unit. We calculated vehicle speed and number plate recognition using the following processes.

Generally, a license plate detection system has to solve two problems: where a license plate is located and how big it is (Howell, 2015). Usually, the candidate position of characters in the license plate is first identified, and the bounding box of the license plate is determined later. There are many challenges in license plate detection in an open environment, such as various observation angles from cameras, background clutter, differently sized license plates, poor image quality from uneven lighting conditions, and multi-plate detection. In terms of drone traffic surveillance in Tonga, the increase in the number of road users today has led to evaluating the use of technologies for monitoring and control. However, the environment posed extra challenges for instance, the

existence of tree branches over public roads, low overhead phone and power lines and tall coconut trees near roads. These can be major obstruction as they block the drone's view and hide vital evidence.

The Quad-copter that was used to take images for data collection purpose in this study used ISO-100, shutter 6400, EV-1 and Fnum-F2.8 camera settings. Figure 1 specifies the steps in image processing that we used. Step one deals with image acquisition, two deals with pre-processing processes which concerns with ways that increases the chances for success. Step three deals with partitioning of the image into parts or objects and step four concerns with the conversion of the input data into a form that is suitable for computer processing. Step five deals with the description of the image. This is concerning the extraction of the features that can be used to differentiate one object from the other. Step six deals with recognition of the image, assigning a label to an object based on its descriptors. Final step is concerned with the interpretation of the image to assign meaning to the recognised objects. We developed this model in the research and applied it to the actual processing of numberplates. The number plate processing required morphological processes, a conversion of the colour imagery from the drone camera to greyscale is, and then a cropping and a dilation of the image so that only the number plate could be fed through the character recognition algorithm. With the limited number (6) of characters on a number plate processing was very quick. When the system is connected to a reference database then immediate data match can be achieved to find the owner of the conformance contract.

The speak calculations could proceed in two fashions. The simplest way is to take time elapses as vehicles move between measured points. However the resultant is an average speed and the distance between the markers has to be sufficient that the entry image and exit image are distinct, and that the margins for visual error calculations are also factored. Consequently we opted for pixel by pixel image processing to calculate instantaneous speeds. Drone traffic surveillance provides many advantages for instance; low cost, easy to deploy, high mobility, large view scope, uniform scale, and so on (Goradia & Xi, 2012; The Economist, 2015). The drone can adjust its flying altitude in order to capture the target object from various angles in order to have the best image. However, data captured by drones contains much more complex information than traditional surveillance video. Footages captured by drone contain not only the traditional data such as traffic flow information but also 360º level data for each vehicle such as vehicle's trajectory data, lane change data and the cars around data on the road. As a result, extracting a moving object from video frames is a challenge.

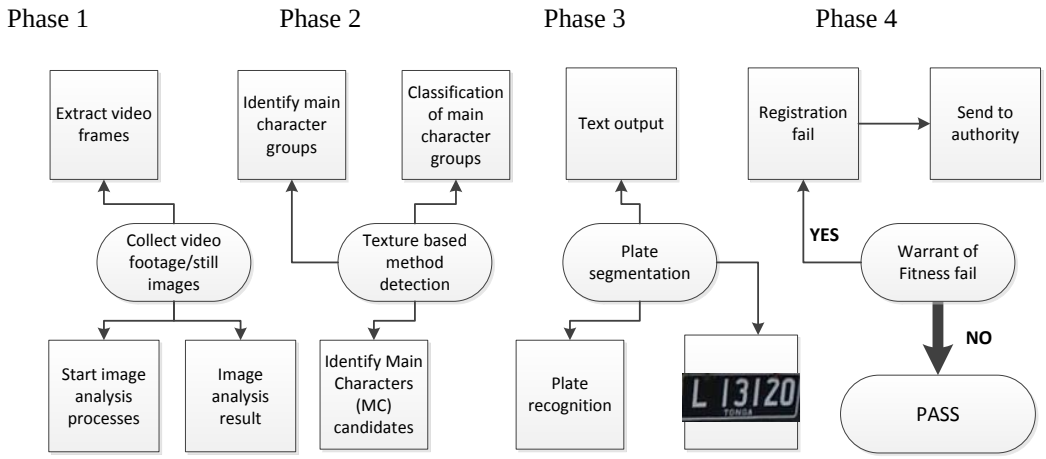


Figure 1: Four phases of image processing employed in this study.

Each vehicle's speed includes two directions – longitudinal and lateral. Usually the lateral value can be assumed to be 0 unless the vehicle makes a lane change. If that happens, the value will change accordingly but relatively small compared to the longitudinal speed. In terms of footage from the drone, any movements in the lateral direction can be the drone's camera motion. In order to minimize this type of errors, only the velocity of vehicle along the road longitudinal direction is considered, which can be expressed by

$$V_{lon} \; {}^{1}/_{4} \, V_x \cos\vartheta + V_y \sin\vartheta$$

It is possible to acquire each pixel's velocity however, due to the variability nature of the optical flow at pixel level; the accuracy of the vehicle's speed obtained is questionable. As a result, the optical flow is used to detect the shape of each vehicle, and the value of V_{lon} is converted into a binary value in the following equation where τ and σ are the maximum and minimum rational speeds.

$$F_{OF}(i,j) = \begin{cases} 1 V_{Ion} \geq \tau \text{ and } V_{Ion} \leq \sigma \\ 0 \text{ otherwise} \end{cases}$$

This relates to the drone's altitude and its stability when the image is captured.

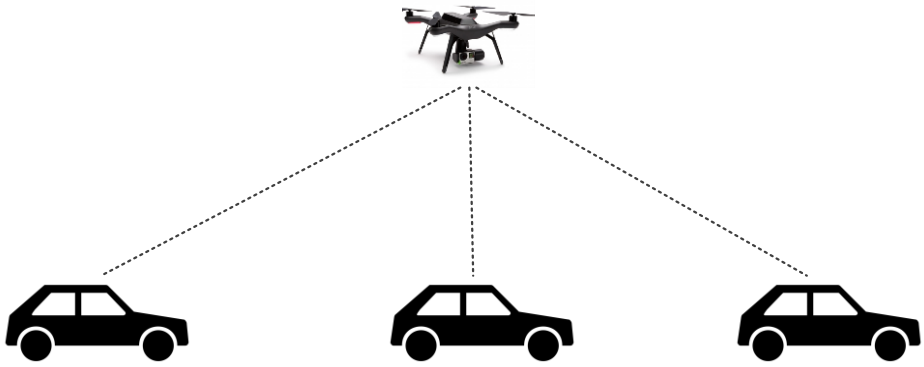


Figure 2: speed calculation images

Drones have proven to be beneficial, effective and efficient in observing and surveilling activities on the ground (Kovar, 2015; Paganini, 2015). However, calculating speed of moving vehicles from a moving camera such as those on drones in real-time is a difficult task due to the potential variability in moving frames, movement and vibration of the drone's camera in the air caused by wind force, the vehicle will change its position on the video as it is travelling so the angle will change accordingly. The settings of the drone's camera might also change such as its brightness values and colours. As a result, stationary and moving cameras call for different processing approaches. In our research we produce the computational algorithms for both instances but only did the calculations and practice for a stable drone. Hence, when the image from the drone's footage is acquired, the point of interest is characterised by its vector features composed of 128 values. The feature matching process started continuously in order to detect changing similarity points and to filter the most stable estimate. The filtering process delivers two images of the vehicle that are very close in time, and hence the distance covered by each image object can be simply computed.

BIG DATA PROBLEMS

The current technologies used for traffic surveillance seem to be inadequate, inasmuch as, purposely placed cameras can only photograph speeding vehicles or vehicles proceeding through red lights. Such images cannot give investigators information such as behaviour of the driver that leads to a traffic accident. For instance, were they weaving all over the road even though they were driving slowly? This type of question needs a different type of monitoring system. However, although there are many benefits of such system, the main challenge is managing big data issues (Rowlingson, 2004; Pooe et al., 2012). The small digital cameras are very efficient and they produce huge amounts of image and other metric data. Big data is a term used to describe sets of electronic data both structured and unstructured, that are large and complex. The current data processing techniques or methods, database applications and software are insufficient. The main issues of big data include capturing, searching, storage, analysis, query and updating and managing information privacy just to name a few. If this is the case then the question is, how big is big data? An example of big data might be Petabytes (1,024 terabytes) or Exabyte (1,024 petabytes) of data consisting of billions to trillions of records. From the data that was collected for this study, a 60 seconds of footage from the quad-copter equals to 446.25MB of video data, this was taken with ISO-100, shutter 6400, EV-1 and Fnum-F2.8 at 25 frames per second, a rate of 60Mbps. At this rate, a drone that can take 30 minutes of video, that is 446.25 x 30 = 13 Gigabytes and 387.5 Megabytes of data. This amount of data will multiply by the number of hours worked each day, and by the number of drones being used by the police department. In terms of storage, if the traffic surveillance system in Tonga employs five quad-copters, video footage streaming from five drones will be 133 Gigabytes and 875 Megabytes of data. Even for small to medium amounts of data the speed of extraction and access to specific segments of data can be hard to manage when the volumes get great. If the traffic surveillance system in Tonga employed five quad-copters, and they observe traffic behaviours every day but only at peak hours, 3 hours in the morning and 3 hours in the evening. Capturing six hours a day of streaming video data from five drones, that's 823 Gigabytes 250 Megabytes of video data. If 823,250 is multiplied by 365 days, that's over 293 Terabytes of data in one year.

A FORENSICALLY READY DRONE INFORMATION SYSTEM

The notion of digital forensic readiness is to meet the objectives for a system that is used in a digital investigation to maximise its ability to collect reliable evidence while minimising the cost (Beebe et al., 2005; Paganini, 2014). The forensic readiness objectives are designed to:

a) To gather admissible evidence legally and without interfering with business processes.
b) To gather evidence targeting the potential crimes and disputes that may adversely impact an organisation.
c) To allow an investigation to proceed at a cost in proportion to the incident.
d) To minimise interruption to the business from any investigation.
e) To ensure that evidence makes a positive impact on the outcome of any legal action.

The advantages of forensic readiness planning include:

Preparing for potential need for digital evidence. If an organisation has to go to litigation and if digital evidence is required, electronic evidence is required to be collected and stored in an appropriate manner so that it is readily available when requested and can be admissible in the court of law. In this process, incident response, disaster recovery and business continuity policies are improved by a forensic readiness policy or plan.

Minimising the cost of the investigation. In the case of a drone surveillance system, the evidence is already gathered, the investigator needs only to process, analyse and review the video footages. This enables faster and more efficient investigation process as a result, the cost of the investigation is minimised and interruption to businesses' normal operation is also minimised. Figure 3 illustrates the systems process architecture.

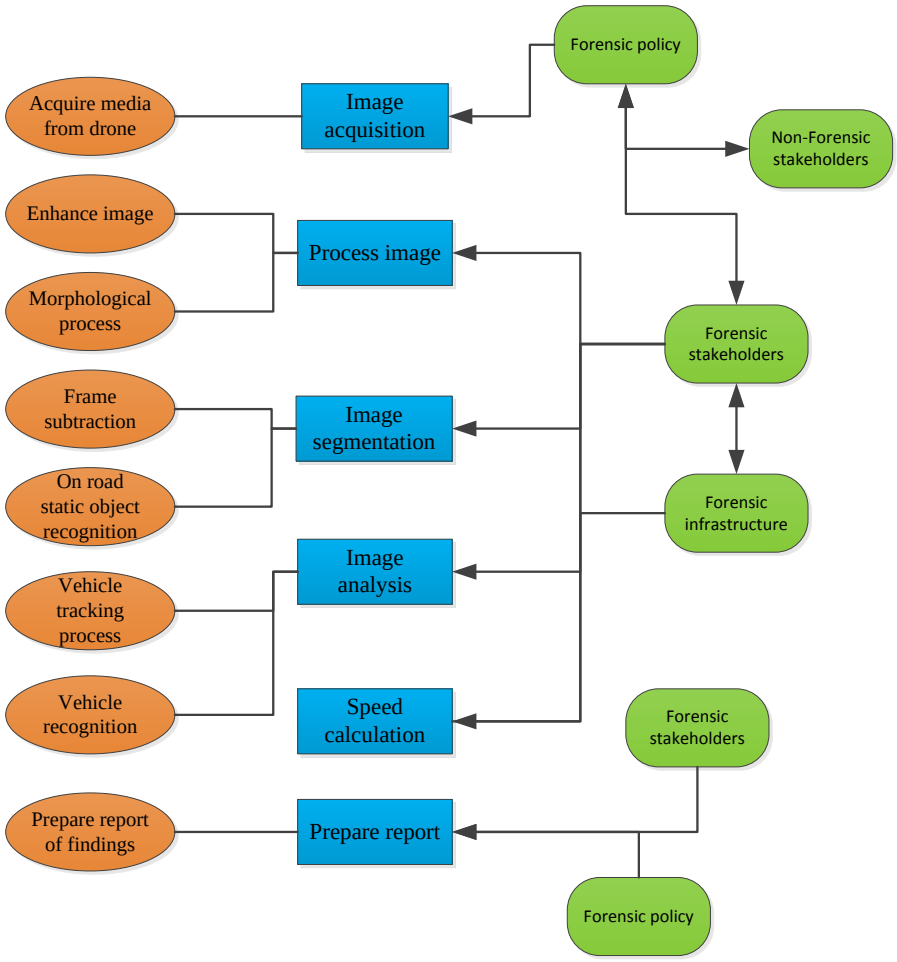


Figure 3: A forensic readiness processes for drone traffic surveillance information system

The requirement to process big data from a drone surveillance program for traffic conformance can be mitigated by an efficient and effective information systems architecture. We propose a structure that is supported by both technology and personnel that can efficiently and effectively mine the mountain of data delivered by both exception and by archival extraction. Real-time data is most efficiently processed by exception filters and this

data can be rapidly transmitted to enforcement officers for further follow-up. Stored data requires both storage and mining capabilities. We propose in figure 4 an information systems architecture that maps operational and infrastructure readiness onto the response units liable to act on the information. The information itself has been categorised before extraction and prepared by due processes that are compliant by law and by policy for its usage. Such a system is a necessary addition to any new technologies that may be used for traffic control and conformance.

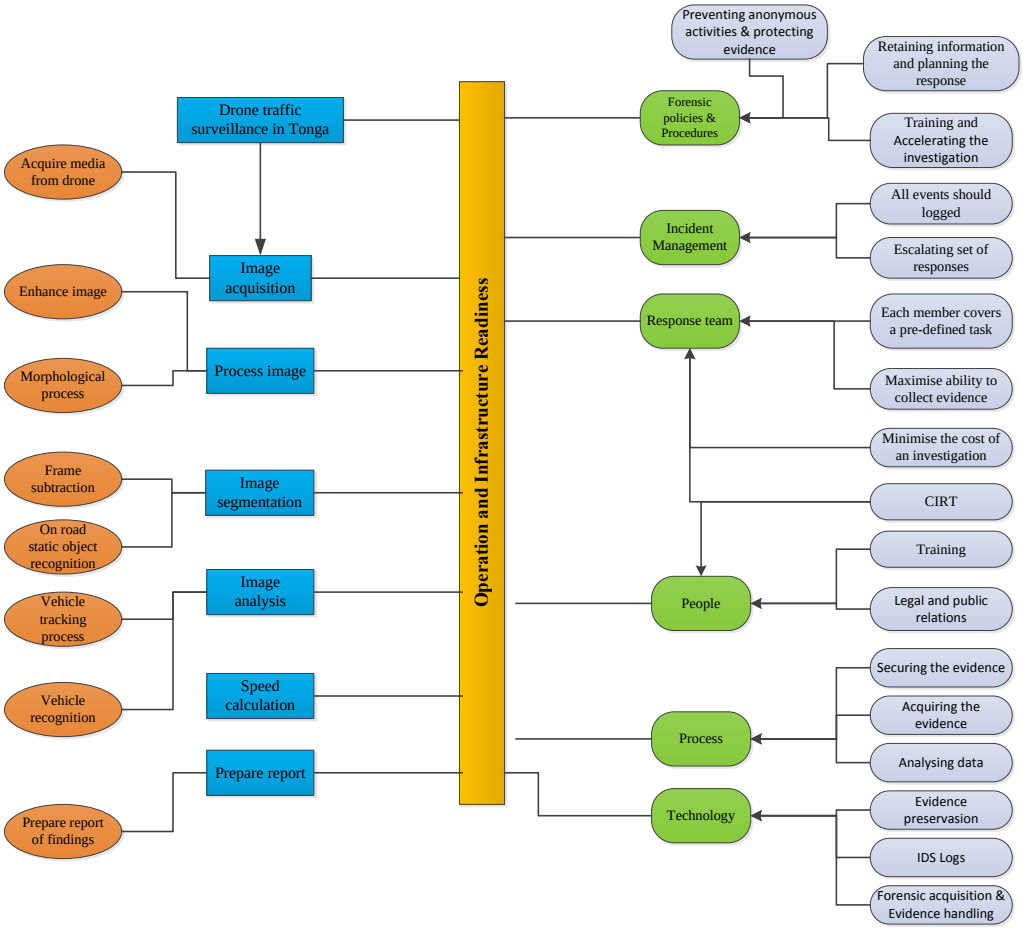


Figure 4: Information Systems architecture for drone data

CONCLUSION

The study was aimed to fill the identified gap in the literature and in practice by developing an information system that is forensically ready to use for drone traffic surveillance in Tonga. The relatively small size of each island and the intensity of traffic allowed lightweight and economical quad copters to be used in this research. The proposal is innovative and as a prototype has added value to current traffic monitoring and control systems. Further research is proposed into the efficient utilisation of the resource so that effective monitoring of vehicle behaviour can be achieved within stringent economic constraints for data. The systems reviewed and shown in this paper provide a cost-effective solution to the constraints imposed by the problem context.

REFERENCES

Ag Ackerman, S. (2013).Welcome to the age of big drone data. Retrieved August 14, 2016, from http://www.wired.com/2013/04/drone-sensors-big-data/

Ahmed, D. T., & Hossain, M. A. (2012). Dynamic prioritization of multi-sensor feeds for resource limited surveillance systems. Proceedings of the 2012 IEEE Conference on Instrumentation and Measurement Technology(I2MTC) (pp.412-416). Graz: IEEE.

Amran, A. R., Saad, A., &AbdRazak, M. R. (2014).An evidential network forensics analysis with metrics for conviction evidence.Proceedings of the 2014 4th International Conference on the Engineering Technology and Technopreneuship (ICE2T) (pp. 73-78). Kuala Lumpur: IEEE.

Bakhtari, A., Naish, M. D., Eskandari, M., Croft, E. A., &Benhabib, B. (2006).Active-vision-based multisensor surveillance - an implementation.IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 36(5), 668-680.

Beebe, N. L., & Clark, J. G. (2005).A hierarchical, objectives-based framework for the digital investigations process.Digital Investigation, 2(2), 147-167.

CambraBaseca, C., Diaz, J. R., &Lloret, J. (2013). Communication Ad Hoc Protocol for Intelligent Video Sensing Using AR Drones. Proceedings of the 2013 IEEE Ninth International Conference on the Mobile Ad-hoc and Sensor Networks (MSN) (pp. 449-453). Dalian: IEEE.

Cheng, H. Y., Weng, C. C., & Chen, Y. Y. (2012).Vehicle Detection in Aerial Surveillance Using Dynamic Bayesian Networks.IEEE Transactions on Image Processing, 21(4), 2152-2159.

Cusack, B. & Khaleghparast, R. (2005). Evaluating Small Drone Surveillance Capabilities to Enhance Traffic Conformance Intelligence. Proceedings Australian Security Conference, SRI, Perth , Australia.

Dijk, J., van Eekeren, A. W., Rojas, O. R., Burghouts, G. J., &Schutte, K. (2013). Image processing in aerial surveillance and reconnaissance: from pixels to understanding. Proceedings of the SPIE Security and Defence International Society for Optics and Photonics (pp.1-17). Netherlands: SPIE.

Goradia, A. and Xi, N. (2012).Modeling and design of mobile surveillance networks using mutational analysis approach. IEEE Transactions on Cybernetics , 42(2), 342-358.

Howell, E. (2015). What is a Drone? Retrieved July 25, 2016, from http://www.space.com/29544-what-is-a-drone.html

Kastrinaki, V., Zervakis, M., &Kalaitzakis, K. (2003).A survey of video processing techniques for traffic applications.Image and Vision Computing, 21(4), 359-381.

Kovar, D. (2015). Drone Forensics – An Overview. Integriography: A Journal of Broken Locks, Ethics, and Computer Forensics, 1(4), 4-7.

Paganini, P. (2015). Hacking drones: Overview of the main threats Introduction,1(1), 1-17.

Paganini, P. (2014). Privacy and security issues for the usage of civil drones. Retrieved August 2, 2016, from http://resources.infosecinstitute.com/privacy-security-issues-usage-civil-drones/

Pooe, A., &Labuschagne, L. (2012).A conceptual model for digital forensic readiness. Proceedings of the 2012 conference on Information Security for South Africa (ISSA) (pp.1-8). Johannesburg: IEEE.

Rowlingson, R. (2004). A ten step process for forensic readiness. International Journal of Digital Evidence, 2(3), 1-28.

The Economist. (2015). Welcome to the Drone Age: Miniature, pilotless aircraft are on the verge of becoming commonplace. The Economist Newspaper Limited, 1(1), -1.

Thompson, R. M. (2012). Drones in domestic surveillance operations: Fourth amendment implications and legislative responses. Congressional Research Service: Library of Congress, 1(1), 1-24.

SURVEY ON REMNANT DATA RESEARCH: THE ARTEFACTS RECOVERED AND THE IMPLICATIONS IN A CYBER SECURITY CONSCIOUS WORLD

Michael James[1], Patryk Szewczyk[2]
[1]Department of Defence, Australia
[2]School of Science, Edith Cowan University, Perth, Australia

Abstract

The prevalence of remnant data in second hand storage media is well documented. Since 2004 there have been ten separate papers released through Edith Cowan University alone. Despite numerous government agencies providing advice on securing personal and corporate information, and news articles highlighting the need for data security, the availability of personal and confidential data on second hand storage devices is continuing, indicating a systemic laissez faire attitude to data security, even in our supposedly cyber security conscious world. The research continues, but there seems to be a lack of correlation of these studies to identify trends or common themes amongst the results. The fact that this type of research continues to be conducted highlights the deficiencies in the methods used to advertise warnings publicised by Government departments and industry experts. Major media organisations seem reluctant to broadcast these warnings, unless there is a bigger story behind the issue. This paper highlights the ongoing issues and provides insight to the factors contributing to this growing trend.

Keywords
Digital storage devices, remnant data, data recovery, privacy, data sanitisation.

INTRODUCTION

Remnant data on second hand devices has been a topic of research since 1996 (Gutmann, 1996). Over the last 10 years' numerous research projects have been completed on this very topic, but to date there has been no collective quantitative or statistical evaluation of all the results, to establish the existence of any trends or common themes. Some research has been conducted in consecutive years by the same academics, e.g. through Edith Cowan University in 2011, 2012 and 2013 (Szewczyk, Robins, & Sansurooah, 2013; Szewczyk & Sansurooah, 2011; Szewczyk & Sansurooah, 2012), and while these papers have compared the types and sizes of the memory cards purchased, there was no quantitative or qualitative analysis of the data retrieved. Another paper did make some comparisons with six years of research (Jones, Valli, Dardick, Sutherland, & Dabibi, 2009), but was limited to their own research.

Digital Storage

Large volume digital storage media is in great demand and the global market as a whole is predicted to reach $6.2 billion by the year 2022 (ReportBuyer, 2015). Advances in computer technology have lead the world to embrace the concept of big data, where more storage is required. Governments and businesses, have unique storage requirements for daily activities and are turning towards the cloud to meet these growing needs. Eventually as storage devices become obsolete and if still serviceable, will still have a monetary value.

2015 saw a decline of -9.2% in the worldwide personal and entry-level hard disk storage market (Li, 2016). This highlights a possible need for the second hand market for the selling of cheap mechanical storage to meet the needs of the home based user. A search for "used" hard disk drives on the popular online auction site eBay yielded more than 33,000 listings. Other storage media types are not immune to this. New and used thumb drives can differ in price by as much as 900% (eBay, 2016). This type of division of cost can be found across all the different types of storage from flash memory to memory cards to solid state hard drives (SSD).

Performance, larger capacities and cost are the driving factors in the growth of flash storage focused technologies (Bez & Pirovano, 2014). Flash storage technologies were predicted to increase in size during 2016 (Sliwa, 2016), this is corroborated by the announcement that Samsung will release a 32TB SSD in 2017 (Shah, 2016). However, while the cost of SSD Hard Drives remain high, the demand for traditional mechanical drives will remain (Bez & Pirovano, 2014).

Previous Research

Edith Cowan University (ECU) has hosted the Australian Digital Forensics Conference since 2006 (ECU, 2016), where numerous papers from around the world that detail research into recovered remnant data on second hand storage media have been presented. The following papers, some that where presented at the conference, and others that were not, have been analysed in this study:

- I know what you did last summer… An Investigation into Remnant Data on USB Storage Devices Sold in Australia in 2015 (Robins, Williams, & Sansurooah, 2015).
- Information Security Leakage: A Forensic Analysis of USB Storage Disks (Adam & Clarke, 2014).
- Analysis of Deletion Habits On Used USB Thumb Drives (Farden & Diesburg, 2014).
- Sellers Continue to Give Away Confidential Information on Second Hand Memory Cards Sold in Australia (Szewczyk et al., 2013a).
- The 2012 Analysis of Information Remaining on Computer Hard Disks offered for Sale on the Second Hand Market in the UAE (Jones, MArtin, & Alzaabi, 2012).
- The 2012 Investigation into Remnant Data on Second Hand Memory Cards Sold in Australia (Szewczyk & Sansurooah, 2012b).
- A 2011 investigation into remnant data on second hand memory cards sold in Australia (Szewczyk & Sansurooah, 2011).
- Data Remanence in New Zealand: 2011(Roberts & Wolfe, 2011).
- The 2009 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market (Jones, Valli, Dardick, et al., 2009).
- The 2009 Analysis of Information Remaining on USB Storage Devices Offered for Sale on the Second Hand Market (Jones, Valli, & Dabibi, 2009).
- The 2008 Australian study of remnant data contained on 2nd hand hard disks: the saga continues (Valli & Woodward, 2008).
- Who is Reading the Data on Your Old Computer (Mee, 2008).
- An Evaluation of Personal Health Information Remnants in Second Hand Personal Computer Disk Drives (Emam, Neri, & Jonker, 2007).
- Oops they did it again: The 2007 Australian study of remnant data contained on 2nd hand hard disks (Valli & Woodward, 2007).
- An empirical methodology derived from the analysis of information remaining on second hand hard disks (Fragkos, Mee, Xynos, & Angelopoulou, 2006).

Statistical Analysis

All of the research papers where reviewed and the data was collated. As shown in Figure 1 of the fifteen research papers, eleven have resulted in data retrieval from more than half of the items purchased, with 40% of the research resulting in data retrieval from over three quarters of the items purchased. It is interesting to note that of all the research, 11 papers recorded attempts to sanitise devices, either by delete, format or repartitioning the volumes. Of these research projects, seven (64%) recorded a sanitisation attempt rate of less than half of the items purchased, and five (45%) are below one quarter.

The quantity of sensitive personal data recovered is also seen as significant (Figure 2), when you consider this as potentially facilitating the means to engage in identity fraud. Corporate data did not fare any better and it is concerning to note that there exist companies, including those in the ICT industry that do not have effective sanitisation and disposal policies to protect their most valuable commodity – their data.

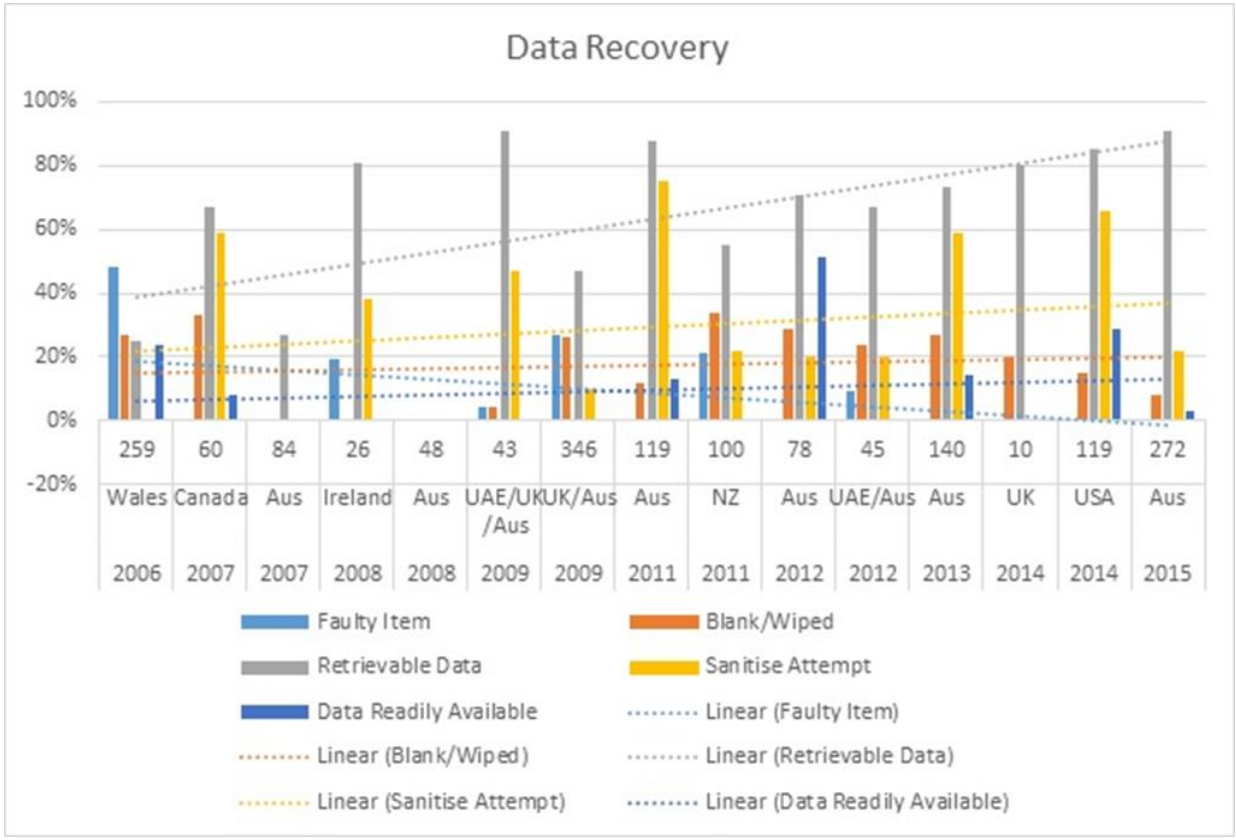


Figure 1 Statistical analysis of 15 remnant data papers with the X axis shows the year, country and number of items purchased

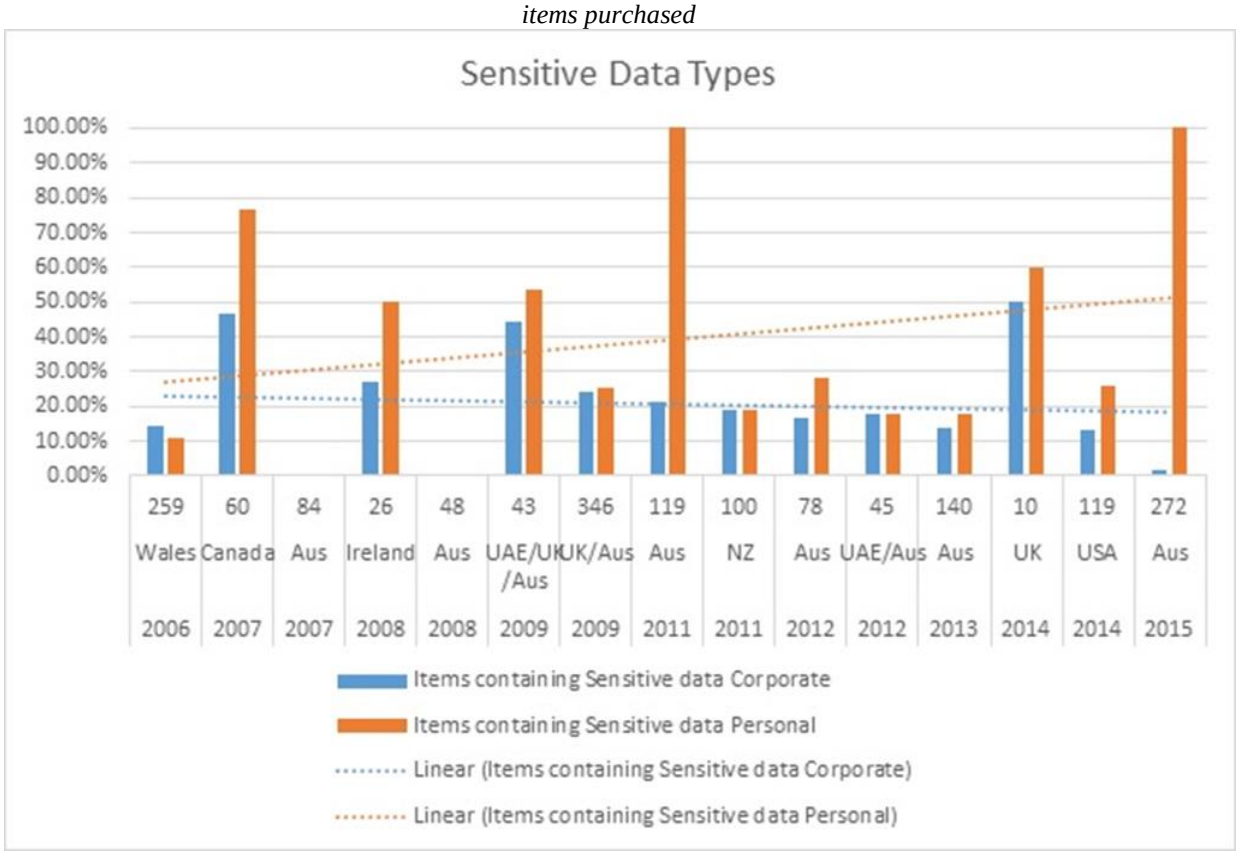


Figure 2 Statistical analysis of 15 remnant data papers with regards to sensitive data types

QUALITATIVE ANALYSIS

Common Themes

All of the research papers detailed very similar outcomes in relation to the sensitive, personal and corporate information that is more than sufficient for identity theft or industrial espionage. The majority of the research papers offered some form of reasoning behind the continuing problem from organisations lacking adequate policies, procedures and strategies that deal with the sanitisation of digital storage to the time consuming nature of developing and implementing not just the strategies and policies but the education of employees.

One common theme across the research is the premise that sellers, both corporate and private do not have an adequate understanding of how 'delete' and 'format' actually work, and have little understanding of how to correctly sanitise their digital media. In 2015 Robins, Williams and Sansurooah postulated that conflicting and misleading advice currently available on the Internet as one possible cause (Robins et al., 2015). However, the fact remains that not enough is known by the general public on how to sanitise digital storage media. Figure 2 shows a steady decline in the amount of corporate data discovered over the 10 years surveyed. This is possibly due to a greater understanding of the problems that exist, however, a number of the research papers show that there are still companies that do not take action, as stated in the 2010 Jones, Valli, Dardick, Sutherland and Dabibi paper.

Corporations spend vast amounts on the protection of data in transit and at rest, but when it comes to the storage device, there seems to be a lax approach. Considering the remnant data that potentially exists, it is strange that some sellers advertise the item for sale as coming from a corporate or government organisation. For those who deal in industrial secrets, this type of hardware would be a goldmine. Take for instance data recovered from a mining company in Australia during 2008. This project recovered supervisory control and data acquisition (SCADA) master plans and passwords, including the Administrator password, as well as photos of the mine site and employees (Valli & Woodward, 2008). Or the data recovered in Australian during the 2015 research, that included legal documents, contracts, agreements and offshore gas transportation documents marked "private and confidential" (Robins et al., 2015). Some companies have taken steps to ensure sanitisation, but are relying on third party organisations to complete the task. The risk here is the possibility that some work is substandard and leaves remnant data behind.

The datatypes that are being recovered are consistent across the entire 10 years of surveyed research. The papers detail the recovery of scanned images of passports, travel itineraries and boarding passes. PayPal and banking data, including login credentials and scanned images of credit cards have been discovered in many of the projects undertaken. Couple this with the ability to identify sellers from photographs recovered from devices and you could have the potential for disastrous consequences to occur for the seller.

Overall a common theme is that all papers report their results to be consistent with similar studies in previous years. The 2011 research completed in New Zealand found their results to be consistent with international research dating back to 2005 (Roberts & Wolfe, 2011), and a 2008 Irish paper reported results consistent with other worldwide research and showed that it was not only a problem in Ireland, but the rest of the world (Mee, 2008). Six years of research with similar results, this shows that education of the masses is either grossly ineffective or non-existent.

Trends

The statistical analysis shows that overall the amount of data retrieved has increased significantly, even though the percentage of data that is seen as 'readily available' has remained relatively static throughout the research (Figure 1). Instances of attempts to sanitise devices has steadily increased, which could be part of lack of knowledge on how 'delete' and 'format' work (this data includes successful sensitisations). An interesting trend is the sharp reduction in the faulty items being sold, which was also reported in the 2009 United Kingdom Australia paper (Jones, Valli, Dardick, et al., 2009).

The trends shown in Figure 2 is that the amount of retrievable data of a personal nature has been increasing whilst corporate data is steadily decreasing. Organisations seem to be becoming more aware of implications and risks associated with disposing of digital media. Eight research papers stated that documents relating to Government organisations were located at some point in the analysis. The latest such find was in 2015 where the construction plans for an Australian army barracks where located. This fact is surprising, when you consider the requirements of the Australian Privacy Principles imposed on all Australian Government Departments from

2013 (OAIC, 2013) and the advice published by the Australian Signals Directorate in their data sanitisation guide (ASD, 2016b).

One trend shows a degree of laziness on the part of sellers, where they are requesting the buyer to remove their data. In 2012 Australian research there were 19 instances where a note requesting the buyer to delete the contents of the device was included with the product (Szewczyk & Sansurooah, 2012). This trend continued in the 2013 Australian research, where a seller supplied a note reporting attempting to remove the data, but was unsure if it was successful and some sellers advertising on the auction listing that devices where being "sold as is" and data would not be deleted due to time constraints (Szewczyk et al., 2013a).

Each year that the research is conducted, the average storage size of the purchased second hand items has increased, this is especially true for the thumb drives and flash memory. There are two possible answers to this, one being that the cost per gigabyte is reducing and the other possibility is that users are requiring more storage space. With the increase in storage capacity, there has been an increase in the amount of retrievable data (Jones, Valli, Dardick, et al., 2009; Robins et al., 2015).

Search Engine Query – "How to delete data"

There is a significant quantity of remnant data being retrieved from second hand devices, and the Robins, Williams and Sansurooah paper mentioned the conflicting and misleading advice on the Internet. As a result an analysis of available information found on the Internet was undertaken for erasing data. The searches were conducted using Google, Bing, Yahoo and Ask, the top four English language search engines (eBiz, 2016), and the browser used was Firefox. The search term "how to erase data from a thumb drive" was used, with only the first page of search results reviewed. To ensure a clinical response, a generic new user account was created for each search engine used and no web accounts logged into.

Table 1 – Search Engine Queries for Data Erasure

	Google	Bing	Yahoo	ASK
Delete/Format	6 Windows 1 Mac	5 Windows	5 Windows	4 Windows
Wipe/Erase	1 Windows 1 Mac	1 Windows 1 Mac	2 Windows 2 Mac	3 Windows 2 Mac
Fill up and delete	1 Windows	1 Windows	1 Windows	1 Windows
Bad links		2		
Advertisements			5	10

Google
Ten results were returned. The first two results were YouTube videos on how to perform a quick format on a flash drive. The video titles were:

- "How to Erase/Delete All Data from Flash Drive Tutorial" (Rivera, 2011)
- "How to erase all content on USB Flashdrive" (Anthoct104, 2011)

Two links took the searcher to sites that detailed how to use the erase a thumb drive, one for the Apple OSX operating system and the other for Microsoft Windows. The Windows environment the tools suggested where Disk Wipe and CCleaner and MAC users where advised to open Disk Utilities and select "Erase". One forum site suggested the user fill the drive with junk, then delete the junk and repeat two more times. Of the remaining six results, the sites the user was directed to varied from forum posts to blog style instructions on how to perform a delete or format of the device. The interesting result is the descriptions used across the results, these ranged from "how to format" to "how to clear" to "how to delete".

Bing
Ten result were returned. The first two results were links to a web page and a YouTube video. The result titles where:

- "Deleting files in your flash drive or memory card using a PC" (SanDisk, 2008)
- "How to Erase/Delete All Data from Flash Drive Tutorial" (Rivera, 2011)

While there was only one result returned for Apple OSX based computers, it also detailed the use of the OSX Disk Utilities Erase function. Of the Windows based results, two where links that displayed content errors and all of the others had titles that ranged from "permanently delete files" to "How to erase all content" to "How to erase a flash drive". After result number five, Bing displayed a link tilted "Videos of how to erase data from a thumb drive". It was observed that this link title related directly to the search criteria, and when opened, a large number of videos where displayed, again with varying titles that contained the words erase, delete, wipe or format.

Yahoo

Ten results were returned, along with 5 advertisements for products, reports and tips relating to wiping data from a device. The first result was a link to a 2011 YouTube video on how to format a USB drive. The first link on how to erase a device was for an Apple OSX at result number 3 and for a PC at result number 4. Result numbers 8 – 10 also described methods for erasing data, including filling the device up with junk and deleting. Yahoo provides featured content labelled as "How to erase data from a thumb drive - Yahoo Answers results". The three items displayed had nothing to do with erasing data from a thumb drive.

ASK

Ten results – the first two results matched those of the searches completed using Google, these being:

- "How to Erase/Delete All Data from Flash Drive Tutorial" (Rivera, 2011)
- "How to erase all content on USB Flashdrive" (Anthoct104, 2011)

10 advertisements where displayed, five above and below the web results with all advertisements at the top repeated at the bottom. Two advertisements where for products that could erase storage media, however, three advertisements where for the recovery of data. Unlike the previous search engines, Ask.com returned six of the ten results as relating to the erasure of data, the first of these results being at number 4 on the page.

Combined Results

The search engines returned similar content, considering that the 2011 YouTube video is consistently within the top two results across all platforms.

Table 2- Common Search Engine Results for Data Erasure

Result Title	**Position in Results List**			
	Google	**Bing**	**Yahoo**	**Ask**
How to Erase/Delete All Data from Flash drive Tutorial	1	2	1	1
How to erase all content on USB Flashdrive	2	7		2
Deleting files in your flash drive or memory card using a PC	3	1	3	3
how to clear data from a USB stick on a Mac	4		2	
How to Completely Erase a Memory Stick	5		4	4
How to Delete the Files on a USB Flash Drive	6	5	6	6
Permanently delete files from a flash drive	7		9	8
USB - How do I format my USB Flash Drive on a Mac?	8		8	
data leakage - How do I securely erase USB flash drives	9			10
How do I erase or wipe an old Flash Drive?	10	10	7	7
How can I securely erase files from a USB drive		9		9
How to Delete Everything on a USB Flash Drive			5	5

A large number of search results appear on at least two of the search engines (Table 2). The array of titles suggests information about erasure of data, but unsuspecting users are supplied conflicting information. For example, "How to Erase/Delete All Data from Flash drive Tutorial" is misleading, as the content is about how to format a flash drive and this result is returned as the top option across all platforms. Of concern is that it was posted in 2011 and this poses the question around how and why search engines return particular results.

DISCUSSION

The world as we know it is in the grip of cybercrime epidemic, in 2015 Steve Morgan, writing for Forbes, stated "Cyber-attacks are costing businesses $400 - $500 billion per year" (Morgan, 2015). In 2016 ABC News reported that Cybercriminals are increasingly targeting Australian consumers (Taha, 2016). The threat of identity theft is real, and this research indicates that people, corporations and governments are taking minimal or no precautions to protect their most valuable asset – their sensitive and identifying data. The research projects analysed shows instances of recovering data that could result in the theft of an identity. The result also showed that while the amount of sensitive corporate data recovered is decreasing, it is still being recovered. Consumers are constantly told how an organisation was hacked or succumbed to an attack, and the incidents of identity theft through the theft of hardcopy mail (Edwards, 2015). But there seems to be very little heard on the discovery of information that can be of value to the identity thief through the recovery of remnant data.

The Australian Signals Directorate has posted information on how to protect data from being divulged (ASD, 2016b), but this document and the subsequent information is difficult to locate. In the 2013 Australian research project it was reported that eBay no longer provide warnings to sellers on the sanitisation of data from storage devices put up for sale (Szewczyk, Robins, & Sansurooah, 2013). These warnings where reported as being provided by eBay in the 2011 Australian paper (Szewczyk & Sansurooah, 2011). This may indicate a complacency exhibited in the attitudes of the sellers. Indeed, some private organisations have conducted their own research into remnant data. Avast antivirus company detailed the purchase of 20 used smartphones from eBay and the personal data that was recovered from these devices (Hořejší, 2014). Corporate and government cloud users have the ability and resources to negotiate the level of control they have over not only the sovereignty of their data, but also the level of oversight on the sanitisation of storage media. The average user does not have such an ability, without paying a significant fee.

CONCLUSION

Unless there is a serious effort to increase the level the education, the issue of recoverable remnant data will continue. As the world moves in the direction of a paperless office, the information that will become available increases exponentially. Many households receive utility bills, bank statements and share dividend statements via email. The world's data holdings stood at 4.4 zettabytes in 2013 and is predicted to grow to 44 zettabytes by 2020 (Khoso, 2016). This large quantity of data has the potential to encompass a significant portion of personal and confidential data that could also end up on the second hand market. This paper has highlighted areas that require further research, these being; an in-depth study of how search engines interpret the needs of the user and then display relevant links; and the quality of literature and supporting information that is supplied to end-users in the public domain.

REFERENCES

Adam, A., & Clarke, N. L. (2014). Information Security Leakage: A Forensic Analysis of USB Storage Disks United Kingdom.

Anthoct104 (Writer). (2011). *How to erase all content on USB Flashdrive*. Retrieved from https://www.youtube.com/watch?v=8sDJ0N1uOuw.

ASD. (2016a). *Cloud Computing Security Considerations. Information Security Advice*. Retrieved from http://www.asd.gov.au/publications/protect/cloud_computing_security_considerations.htm

ASD. (2016b). *Data Spill Sanitisation Guide. Canberra Australia: Department of Defence*. Retrieved from http://www.asd.gov.au/publications/protect/data_spill_sanitisation_guide.htm.

Bez, R., & Pirovano, A. (2014). *Overview of non-volatile memory technology: markets, technologies and trends*. In Y. Nishi (Ed.), Advances in Non-volatile Memory and Storage Technology.

eBay. (2016). *eBay*. Retrieved from http://www.ebay.com.au/

eBiz. (2016). *Top 15 Most Popular Search Engines | September 2016*. Retrieved from http://www.ebizmba.com/articles/search-engines

ECU. (2016). *Australian Digital Forensics Conference*. Retrieved from http://ro.ecu.edu.au/adf/index.2.html

Edwards, M. (2015). *Identity theft: More than 770,000 Australians victims in past year*. ABC News. Retrieved from http://www.abc.net.au/news/2015-04-14/identity-theft-hits-australians-veda/6390570

Emam, K. E., Neri, E., & Jonker, E. (2007). An Evaluation of Personal Health Information Remnants in Second-Hand Personal Computer Disk Drives. *JMIR Publications*, 9(3).

Farden, M. A., & Diesburg, S. (2014). *Analysis Of Deletion Habits On Used USB Thumb Drives*. In U. o. N. Iowa (Ed.).

Fragkos, G., Mee, V., Xynos, K., & Angelopoulou, O. (2006). *An empirical methodology derived from the analysis of information remaining on second hand hard disks.* Paper presented at the Second European Conference on Computer Network Defence.

Gutmann, P. (1996). *Secure Deletion of Data from Magnetic and Solid-State Memory*. Auckland, New Zealand.

Hořejší, J. (2014, 9 July 2014). *How Avast recovered 'erased' data from used Android phones*. Retrieved from https://blog.avast.com/2014/07/09/android-foreniscs-pt-2-how-we-recovered-erased-data/

Jones, A., MArtin, T., & Alzaabi, M. (2012). *The 2012 Analysis of Information Remaining on Computer Hard Disks Offered for Sale on the Second Hand Market in the UAE*. Paper presented at the 10th Australian Digital Forensics Conference, Perth Western Australia.

Jones, A., Valli, C., & Dabibi, G. (2009). *The 2009 Analysis of Information Remaining on USB Storage Devices Offered for Sale on the Second Hand Market*. Paper presented at the 7th Australian Digital Forensic Conference, Perth Western Australia.

Jones, A., Valli, C., Dardick, G. S., Sutherland, I., & Dabibi, G. (2009). *The 2009 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market*. Paper presented at the 8th Australian Digital Forensics Conference, Perth Western Australia.

Khoso, M. (2016, 13 May 2016). *How Much Data is Produced Every Day?* Retrieved from http://www.northeastern.edu/levelblog/2016/05/13/how-much-data-produced-every-day/

Li, J. (2016). *Worldwide Personal & Entry-Level Storage Market Declined in 2015*, According to IDC International Data Corporation, IDC tracker. Retrieved from https://www.idc.com/getdoc.jsp?containerId=prUS41021816

Mee, V. (2008). Who is Reading the Data on Your Old Computer? *Journal of Digital Forensics, Security and Law*, 3(1), 25-34.

Morgan, S. (2015). *The Business of Cybersecurity: 2015 Market Size, Cyber Crime, Employment, and Industry Statistics*. Retrieved from http://www.forbes.com/sites/stevemorgan/2015/10/16/the-business-of-cybersecurity-2015-market-size-cyber-crime-employment-and-industry-statistics/#130eeabe10b2

OAIC. (2013). *Australian Privacy Principles*. Retrieved from https://www.oaic.gov.au/privacy-law/privacy-act/australian-privacy-principles

ReportBuyer. (2015). *Global Digital Storage Devices Market Outlook (2014-2022).* PR Newswire. Retrieved from http://www.prnewswire.com/news-releases/global-digital-storage-devices-market-outlook-2014-2022-300145285.html

Rivera, L. (Writer). (2011). *How to Erase/Delete All Data from Flash drive Tutorial --(Even hidden files!).* Retrieved from https://www.youtube.com/watch?v=8wWB6Ee469I

Roberts, D., & Wolfe, H. B. (2011). *Data remanence in New Zealand: 2011*. Paper presented at the 9th Australian Digital Forensics Conference, Perth, Western Australia.

Robins, N., Williams, P. A. H., & Sansurooah, K. (2015). *I know what you did last summer… An Investigation into Remnant Data on USB Storage Devices Sold in Australia in 2015*. Paper presented at the Australasian Computer Science Week, Canberra Australia.

SanDisk. (2008). *Deleting files in your flash drive or memory card using a PC*. Retrieved from http://kb.sandisk.com/app/answers/detail/a_id/2281/~/deleting-files-in-your-flash-drive-or-memory-card-using-a-pc

Shah, A. (2016). *Samsung's massive 32TB SSD includes cutting-edge 3D chip technology*. Retrieved from http://www.pcworld.com/article/3105875/storage/samsungs-massive-32tb-ssd-includes-cutting-edge-3d-chip-technology.html.

Sliwa, C. (2016). *Flash technologies remain hot in 2016, experts predict*. Retrieved from http://searchsolidstatestorage.techtarget.com/news/4500273061/Flash-technologies-remain-hot-in-2016-expects-predict.

Szewczyk, P., Robins, N., & Sansurooah, K. (2013a). *Sellers Continue to Give Away Confidential Information on Second Hand Memory Cards Sold in Australia*. Paper presented at the 11th Australian Digital Forensics Conference, Perth, Western Australia.

Szewczyk, P., & Sansurooah, K. (2011). *A 2011 investigation into remnant data on second hand memory cards sold in Australia*. Paper presented at the 9th Australian Digital Forensics Conference, Perth, Western Australia.

Szewczyk, P., & Sansurooah, K. (2012). *The 2012 Investigation into Remnant Data on Second Hand Memory Cards Sold in Australia*. Paper presented at the 10th Australian Digital Forensic Conference, Perth, Western Australia

Taha, M. (2016). *Cybercriminals increasingly targeting Australia as a launch pad for cybercrime*. ABC News. Retrieved from http://www.abc.net.au/news/2016-02-26/cyber-criminals-increasingly-targeting-australia/7203478

Valli, C., & Woodward, A. (2007). *Oops they did it again: The 2007 Australian study of remnant data contained on 2nd hand hard disks*. Paper presented at the 5th Australian Digital Forensics Conference, Perth, Western Australia.

Valli, C., & Woodward, A. (2008). *The 2008 Australian study of remnant data contained on 2nd hand hard disks: the saga continues*. Paper presented at the 6th Australian Digital Forensics Conference, Perth, Western Australia.